

University of Groningen

VIOT 2018: Duurzame Taalbeheersing

Jansen, C. J. M.

DOI:
[10.5117/9789463720410](https://doi.org/10.5117/9789463720410)

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2019

[Link to publication in University of Groningen/UMCG research database](#)

Citation for published version (APA):

Jansen, C. J. M. (2019). *VIOT 2018: Duurzame Taalbeheersing: Lezingen van de veertiende VIOT-conferentie gehouden op 17, 18 en 19 januari 2018 aan de Rijksuniversiteit Groningen*. (1 redactie) Amsterdam: Amsterdam University Press. <https://doi.org/10.5117/9789463720410>

Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

2019
VOLUME 41
NO. 1

VIOT 2018:
Duurzame Taalbeheersing
Lezingen van de veertiende VIOT-
conferentie gehouden op 17, 18 en
19 januari 2018 aan de Rijksuniversiteit
Groningen

Onder redactie van Carel Jansen

Tijdschrift voor
Taalbeheersing

VIOT

Amsterdam
University
Press



VIOT 2018: Duurzame Taalbeheersing

VIOT 2018: Duurzame Taalbeheersing

Lezingen van de veertiende VIOT-conferentie gehouden
op 17, 18 en 19 januari 2018 aan de Rijksuniversiteit Groningen

Onder redactie van Carel Jansen

Amsterdam University Press

Cover design: Coördesign, Leiden
Lay-out: Pre Press Media Groep, Zeist

ISBN 978 94 6372 041 0
e-ISBN 978 90 4854 409 7 (pdf)
DOI 10.5117/9789463720410
NUR 623



Creative Commons License CC BY NC ND (<http://creativecommons.org/licenses/by-nc-nd/3.0>)

© The authors / Amsterdam University Press B.V., Amsterdam 2019

Some rights reserved. Without limiting the rights under copyright reserved above, any part of this book may be reproduced, stored in or introduced into a retrieval system, or transmitted, in any form or by any means (electronic, mechanical, photocopying, recording or otherwise).

Inhoud

Voorwoord	9
'Wapper wat meer met je handen' <i>De invloed van gebaren op retentie en sprekerswaardering bij een informatieve presentatie</i> SJAAK BAARS EN BAS ANDEWEG	11
Opvattingen van mbo-docenten over beroepsgerichte taal- en communicatie-eisen en de aansluiting op een dynamische arbeidsmarkt ANNET BRUIN, JACQUELINE VAN KRUININGEN, KEES DE GLOPPER, ARJAN VAN DER MEIJDEN EN NICOLE VAN DER MEULEN	27
Is <i>translanguaging</i> een duurzame strategie voor het hoger onderwijs in Zuid-Afrika? ADELIA CARSTENS	41
En dan zit je met de gebakken peren! <i>Strategisch manoeuvreren met het argumentum ad consequentiam</i> FRANS VAN EEMEREN EN BART GARSSSEN	55
Tekstuele kenmerken van misleidende journalistiek <i>De zaak-Ramesar</i> CHARLOTTE GOVAERT, LUUK LAGERWERF EN CÉLINE KLEMM	69
Strategisch manoeuvreren met stijl <i>Een systematische benadering</i> TON VAN HAAFTEN EN MAARTEN VAN LEEUWEN	83
Probleembesprekingen met samenwerkende kleuters FRANS HIDDINK	97

Wat leert onderzoek naar overtuigende teksten over het ontwerpen van overtuigender teksten?	113
<i>Een overzicht van meta-analytische studies</i>	
HANS HOEKEN	
Zijn concrete argumenten doorslaggevender?	137
<i>Een experiment naar het effect van mate van concreetheid op sociale oordelen</i>	
LETTICA HUSTINX, IRIS HOFSTRA EN ANNE JANSSEN	
De invloed van het standpunt op de beoordeling van <i>ad populum</i> -argumentatie	151
HENRIKE JANSEN	
Pressie en argumentatie in maatschappelijke discussies	167
JAN ALBERT VAN LAAR EN ERIK C. W. KRABBE	
De 'stem van het volk'	181
<i>Populisme en perspectief</i>	
MAARTEN VAN LEEUWEN EN FREEK VAN VLIET	
Het effect van monologen en dialogen in radioreclame	197
FRANK VAN MEURS, BERNA HENDRIKS EN DILEK KÖKSAL	
Een significant probleem	211
GERBEN MULDER	
De rol van semantische afstand in visuele metaforen	223
MARGOT VAN MULKEN, LUUK LAGERWERF EN IRIS BLOKLAND	
De rol van gelijkenis bij narratieven	237
JOËLLE OOMS, JOHN HOEKS EN CAREL JANSEN	
Retorische vragen van de rechter als wrakingsgrond	251
<i>Een argumentatie-analytisch model</i>	
H. JOSÉ PLUG	
De redacteur als tekstbeoordelaar	265
<i>Acquisitie en redactie bij Nederlandse algemene uitgeverijen</i>	
EVERDIEN RIETSTAP	

Een pragma-dialectische reconstructie van de discussiebijdragen van arts en patiënt in 'shared decision-making'	281
A. FRANCISCA SNOECK HENKEMANS EN JEAN H.M. WAGEMANS	
Blijvende impact	295
<i>Onderzoek robuust en relevant maken</i>	
CATHERINE E. SNOW	
Ontnuchterende voorlichtingscampagnes over comazuipen bij studenten	309
<i>Een experimenteel onderzoek naar de overtuigingskracht en effectiviteit van humor en angst, en inspelen op fysieke en sociale gevaren</i>	
SARA VERBRUGGE EN ROMANE MORREALE	
Effecten van anti-rookboodschappen bij jongeren	323
Een conceptueel replicatieonderzoek	
HELEEN VAN WIJK EN CAREL JANSEN	

Voorwoord

Van 17 tot en met 19 januari 2018 vond aan de Rijksuniversiteit Groningen de veertiende VIOT-conferentie plaats, onder de noemer *VIOT 2018*. Tijdens deze driedaagse conferentie van de Vereniging Interuniversitair Overleg Taalbeheersing stond het thema *Duurzame Taalbeheersing* centraal. Duurzaamheid vormde vanuit verschillende invalshoeken de rode, of liever groene draad tijdens deze VIOT-conferentie. In 85 lezingen en negen symposia werd de aandacht gevestigd op de vraag wat in theorievorming, onderzoek en toepassingen op het terrein van de Taalbeheersing en de Communicatiekunde van blijvende waarde is. Ook kwamen methodologische kwesties aan de orde rond herhaalbaarheid van onderzoek en robuustheid van uitkomsten.

De organisatiecommissie (Jacqueline van Kruiningen (voorzitter), Kees de Glopper, John Hoeks en Carel Jansen) dankt de collega's die betrokken waren bij de organisatie van de conferentie voor hun inzet. Van de Groningse afdeling CIW Groningen waren dat Aline Douma, Noortje Hemmen, Annemiek Nijp en Aniek Straatman, en van het Groningen Congres Bureau Mariska Pater en Sharon de Puijsseelaar. Dank gaat ook uit naar de sponsors van de conferentie: het Groninger Universiteitsfonds, het Centre for Language and Cognition Groningen, de Taalunie en het bestuur van de VIOT.

In deze bundel zijn twee van de vier keynote-bijdragen opgenomen. Catherine Snow (Harvard Graduate School of Education) vestigt in haar bijdrage *Blijvende impact* de aandacht op duurzaamheid en relevantie in onderwijsonderzoek. Hans Hoeken (Universiteit Utrecht) presenteert in zijn bijdrage *Wat leert onderzoek naar overtuigende teksten over het ontwerpen van overtuigender teksten?* een overzicht van meta-analytische studies over dit thema. Daarnaast bevat deze bundel twintig bijdragen van lezinghouders uit de parallele sessies.

Dank verdienen de 28 reviewers die de bijdragen van constructief commentaar hebben voorzien. Hun namen staan hieronder. Natuurlijk gaat de dank vooral uit naar de auteurs die hun lezing hebben bewerkt tot een bijdrage aan deze VIOT-bundel. Net als de vorige VIOT-bundels biedt ook

deze bundel daarmee een gevarieerd overzicht van de thema's en onderzoeksmethoden in actueel taalbeheersingsonderzoek – dit keer met bijdragen uit niet alleen België en Nederland, maar ook uit Zuid-Afrika en de Verenigde Staten.

Speciale dank gaat uit naar Aline Douma voor haar ondersteuning van de redactie en voor de vertaling van het artikel van Catherine Snow. Dank gaat ook uit naar Amsterdam University Press (AUP) voor de mogelijkheid om deze bundel als een nummer van het *Tijdschrift voor Taalbeheersing* uit te geven en de bijdragen na publicatie ook in open access-vorm beschikbaar te stellen.

Carel Jansen (redactie)

Reviewers: Bas Andeweg, Huub van den Bergh, Bart Garssen, Myrte Gosen, Hans Hoeken, John Hoeks, Bregje Holleman, Jos Hornikx, Mike Huisjes, Henrike Jansen, Tom Koole, Femke Kramer, Kobie van Krieken, Luuk Lagerwerf, Fons Maes, Els van der Pool, Yfke Ongena, Joëlle Ooms, Albert Oosterhof, Henk Pander Maat, José Plug, Gisela Redeker, Gert Rijlaarsdam, Wilbert Spooren, Leon de Stadler, Ninke Stukker, Dorien Van De Mierop en Sara Verbrugge

'Wapper wat meer met je handen'

De invloed van gebaren op retentie en sprekerswaardering bij een informatieve presentatie

Sjaak Baars en Bas Andeweg
Technische Universiteit Delft

TVT 41 (1): 11–25

DOI: 10.5117/TVT2019.1.001.BAAR

Abstract

'Flap those hands' – The influence of gestures on retention and speaker's assessment in an informative presentation

To convey information during an oral presentation speakers not only use words, they also gesticulate. Their gestures can be divided into iconic, metaphoric, deictic and beat gestures (McNeill, 1992). Beats (repetitive, short movements) are frowned upon by some presentation skills advisers. Earlier research that focused on short speeches, mostly about concrete topological content, found that gestures help the listener to understand and remember the content. Presentation skills courses, however, focus on longer, more abstract informative speeches. To explore how gestures influence both retention and assessment of the speaker in such longer speeches, an experiment was conducted. Participants ($N = 229$) were asked to watch a fifteen minute informative presentation accompanied by PowerPoint slides, either (i) without gestures, (ii) with only beat gestures, or (iii) with a mix of iconic, metaphoric, deictic, and beat gestures. Participants were tested on retention and on their assessment of speaker qualities. An ANOVA showed a significant effect for retention. When a speaker accompanied his speech with only beat gestures, this resulted in higher scores than when this speaker remained in a static position. Also, when the speaker used only beat gestures, he was seen as more 'natural' than when he remained in a static position. These results do not imply causality. They may, however, temper advisers' warnings against using beat gestures in presentations.

Keywords: gestures, beats, retention, oral presentation, education

Inleiding

Bij een informatieve presentatie gaat het primair om de inhoud. Het doel van de presentatie is dat de luisteraar de boodschap begrijpt en onthoudt. Toch maken professionele sprekers zich niet alleen druk om een heldere structuur en een goede formulering. Weervrouw Willemijn Hoebert bijvoorbeeld formuleert het als volgt: ‘Tijdens de presentatie is het belangrijkste dat ik niet afleid van de inhoud. [...] Toen ik begon was ik Floortje Fladder, ik wapperde continu met mijn handen. Daar heb ik met een presentatiecoach aan gewerkt’ (Waarlo, 2017). Het citaat maakt duidelijk dat de spreker ideeën heeft over wat luisteraars wel of niet meenemen als ze luisteren naar haar presentaties. Haar hoofddoel is het overbrengen van de informatie. Ook beginnende sprekers, zoals onze studenten bij het presentatieonderwijs aan de Technische Universiteit Delft, maken zich zorgen over de manier waarop ze overkomen op de mensen in de zaal: moet ik stilstaan, of juist niet? Welke gebaren moet ik maken of juist vermijden?

Auteurs van adviesboeken over mondeling presenteren zien weliswaar de waarde van gebaren in voordrachten, maar ze zijn niet altijd concreet over het soort gebaren dat sprekers zouden moeten hanteren om hun presentatie te optimaliseren. Soms worden weliswaar instructies gegeven om te studeren op gebaren (Kirchner, 1983; Van Eijk, 1987; Van der Meiden, 1991) maar een terugkerende mening is dat sprekersgebaren vooral ‘natuurlijk’ moeten zijn (Blum, 1982; Kirchner, 1983; De Boer, 1986; Mertens, 1992; Wiertzema & Jansen, 2004). Met ‘natuurlijk’ lijkt soms ‘spontaan’ te worden bedoeld, soms ook ‘passend’. Zulke adviezen lijken een leerlingsspreker in een soort spagaat te plaatsen: enerzijds worden gebaren gezien als een gewenste ondersteuning, maar anderzijds moeten ze niet ingestudeerd worden, want dan komen ze gekunsteld over. In dit artikel staat de volgende vraag centraal.

In hoeverre heeft het maken van gebaren tijdens een informatieve presentatie invloed op het publiek?

Hierbij richten we ons op twee aspecten: retentie (een kerndoel bij een informatieve presentatie) en sprekerswaardering, dus: hoe beïnvloeden gebaren het onthouden van de boodschap en hoe beïnvloeden gebaren de waardering van de luisteraars voor de spreker? Uiteraard zijn gebaren slechts een onderdeel van de algemene fysieke gedragingen van een spreker; tijdens een voordracht spelen bijvoorbeeld ook mimiek, intonatie, oogcontact, en het bewegen in een ruimte een belangrijke rol. Het onderzoeken

van deze fysieke gedragingen is een uitdaging, omdat de verschillende aspecten elkaar beïnvloeden: een spreker die enthousiast is over hetgeen hij bespreekt zal dit niet alleen duidelijk maken met zijn intonatie, maar ook met expressie in zijn gezicht (Sanders, 1987). In dit onderzoek richten we ons alleen op de gebaren als mogelijk element in het onderwijs in mondeling presenteren.

Literatuur

Wie spreekt, maakt gebaren. Maar waarom? Helpt het de spreker zelf, of doet hij dit voor de luisteraar? De afgelopen decennia hebben deze vragen de aandacht gehad van onderzoekers uit meerdere disciplines (zie voor een informatief overzicht: Wagner, Malisz, & Kopp, 2014). Krauss en Hadar (1999) bijvoorbeeld veronderstellen dat het maken van gebaren helpt om de juiste woorden te vinden. Opmerkelijk genoeg maken zelfs geboren blinden gebaren als ze een ander iets uitleggen (Iverson & Goldin-Meadow, 1998). Uit onderzoek van Sueyoshi en Hardison (2005) en van Beattie en Shovelton (1999) blijkt dat luisteraars meer onthouden van een boodschap als ze de spreker ook *zien* in plaats van alleen *horen*. Dit resultaat wordt deels toegeschreven aan gebaren, maar hierbij kan ook ander fysiek gedrag een belangrijke rol spelen, zoals mimiek, intonatie en oogcontact. Vanzelfsprekend hebben al deze verschillende gedragingen, eventueel in combinatie, invloed op een voordracht. In dit onderzoek richten we ons echter alleen op effecten van gebaren en beperken we ons tot de vraag of luisteraars profiteren van, of anderszins beïnvloed worden door een gebarende spreker.

Gesticulerende sprekers maken verschillende soorten gebaren (Kendon, 1980). De betekenis van gebaren ligt niet vast en is omgevings- (en vaak ook cultuur-)afhankelijk. Krauss, Morrel-Samuels en Colasante (1991) concluderen uit hun onderzoek zelfs dat gebaren meestal pas hun betekenis krijgen door de begeleidende tekst. Doventaalggebaren vormen hierin een uitzondering. Op basis van eerder onderzoek van Kendon (1980) onderscheidt McNeill (1992) gebaren in vier groepen: (i) iconische gebaren voor een concreet object of concrete actie (bijvoorbeeld: de handen naast elkaar als een open boek), (ii) metaforische gebaren voor abstracte objecten/acties (bijvoorbeeld: een hand over de schouder als je over het verleden praat), (iii) deiktische gebaren als een spreker wijst en (iv) beats: korte, herhalende, simpele bewegingen van de hand of vinger, omhoog of opzij, vaak op het ritme van de zin. McNeill (1992) ziet de waarde van beats in het feit dat deze nadruk kunnen leggen

op woorden. De meeste auteurs benadrukken dat beats geen betekenisvolle informatie overbrengen. In dit artikel maken we daarom onderscheid tussen zogenaamde betekenisvolle gebaren (groep i, ii, en iii) en beats (groep iv).

Maar hebben die verschillende gebaren nu invloed op de manier waarop de luisteraar de informatie verwerkt? In een meta-analyse concludeert Hostetter (2011) dat gebaren een significant, maar beperkt positief effect hebben op een luisteraar. De mate van het effect wordt volgens haar gedomereerd door drie factoren: het soort gebaar (gebaren die verwijzen naar acties zijn effectiever dan gebaren die verwijzen naar abstracte concepten), de leeftijd (kinderen profiteren meer) en de overlap van de gebaren met de gesproken tekst (hoe meer de gebaren informatie bevatten die niet in de tekst is opgenomen, hoe groter het effect). In haar meta-analyse gaat het overigens vooral om deiktische en iconische gebaren. Specifiek onderzoek naar beat-gebaren is minder frequent.

De afgelopen decennia zijn enkele experimenten uitgevoerd waarin nagegaan werd of er verschillen in retentie optraden tussen een situatie waarin een spreker wel of niet gebaarde, en of de spreker meer gewaardeerd werd door het publiek als hij wel of geen gebaren maakte.

Kang, Hallman, Son en Black (2012) probeerden na te gaan of er een effect was voor gebaarstype bij een college van vijf minuten over celbiologie. Er werden drie verschillende video's gemaakt (met (i) betekenisvolle gebaren, met (ii) beats, en (iii) zonder gebaren). Het eerste college bevatte slechts negen gebaren per minuut (totaal 45). Beats werden gedefinieerd als een enkele beweging van de open rechterhand. Proefpersonen bekeken een van deze video's. Kang et al. constateerden geen verschil in retentie tussen de verschillende condities, al scoorde video (i) beter bij vragen die het analytisch vermogen testten. Mogelijk moet de gelijke retentie ook toegeschreven worden aan het feit dat de proefpersonen het college twee keer moesten bekijken, waarbij hen vooraf werd verteld dat ze getest zouden worden.

In een experiment van Maricchiolo, Gnisci, Bonaiuto en Ficca (2009) werd niet naar retentie gekeken maar naar onder meer de sprekerswaardering in combinatie met gebaren. In een video van tweeënhalve minuut betoogde een spreker dat het collegegeld met twintig procent omhoog moest. Er werden verschillende versies voorgelegd die alleen verschillen in het soort gebaren, waaronder (i) iconisch, metaforisch, deiktisch, (ii) beats en (iii) geen gebaren. Proefpersonen evalueerden na afloop onder andere de spreker op categorieën als competentie, warmte en rust. In de versie met betekenisvolle gebaren werd de spreker als competentier beschouwd dan in de versie zonder gebaren.

Behalve in de publicaties over bovenstaande experimenten wordt ook in adviesboeken over mondelinge vaardigheden uitgebreid en uiteenlopend geschreven over gebaren. Schreef in 1940 Ritter nog: 'Het gebaar is een gevaar' (Ritter, 1940, p. 89), latere auteurs als Braas, Kat, Timmer en Ville (2001, p. 66) stellen dat: 'iedereen kan leren om functionele gebaren te maken'. In een corpus van veertig meest populaire adviesboeken in de periode 1980-2010 (Wackers, De Jong, & Andeweg, 2016) zijn er slechts twee auteurs die niets schrijven over gebaren. Dat wil overigens niet zeggen dat het onderwerp door eenieder uitputtend beschreven wordt. Janssen (2002, p. 353) adviseert bijvoorbeeld slechts – zonder verdere precisering – om ondersteunende gebaren te gebruiken. Andere auteurs zijn veel concreter. Kirchner (1983) en Van Eijk (1987) wijden zelfs een volledig hoofdstuk aan handbewegingen. Van der Meiden (1991, p. 162) bijvoorbeeld stelt dat gebaren en tekst congruent moeten zijn: 'geeft u brede lijnen aan, dan mag u de handen langzaam heen en weer, horizontaal laten bewegen. Als u zegt dat er verschillende niveaus in het probleem zitten, moet u lagen aangeven op het horizontale vlak'. Hij waarschuwt voor 'spinaziehakkers' of 'masseurs', mensen die hun beide handen snel op en neer bewegen (p. 80). Krusche (1986, p. 49) stelt dat het gebruik van dergelijke ritmische gebaren aangeeft dat 'de spreker er innerlijk niet bij is'. Steehouder et al. (2006, p. 289) geven aan dat deze beat-bewegingen onbewust gebeuren en 'geen betekenis' hebben. Toch vinden ze deze belangrijk: 'ze verlevendigen uw houding en dragen ertoe bij dat u een gemotiveerde indruk maakt', maar betekenisvolle gebaren worden als beter beschouwd.¹

Regelmatig – door 30 procent van de auteurs in het onderzochte corpus – wordt de nadruk gelegd op 'natuurlijkheid' van de gebaren (onder meer Blum, 1982; Kirchner, 1983; De Boer, 1986; Mertens, 1992; Wiertzema & Jansen, 2004). Blum (1982, p. 105) stelt dat 'aangeleerde gebaartjes' onecht zijn en de spreker ongeloofwaardig maken. Kirchner (1983, p. 54) roept over gebaren: 'Maak er liever geen, dan onechte!' Mertens (1992, p. 14) zegt dat 'toehoorders natuurlijkheid verwachten' [...] 'Iemand die [...] nooit gebaren maakt, moet dat niet willen proberen'. Wat deze natuurlijkheid precies inhoudt, wordt niet duidelijk. Allereerst wordt de suggestie gewekt dat een spreker die zich 'van nature' op een bepaalde manier gedraagt, niet beïnvloed is door oefening en scholing – zijn gedrag is zuiver. Daarnaast lijkt natuurlijkheid ook gebonden te zijn aan spontaniteit – als gedrag dat niet door overwegingen wordt gestuurd. Ten slotte lijkt natuurlijkheid voor de auteurs ook te horen bij passendheid. Krusche (1986, p. 49) gaat er uitgebreid op in als hij stelt dat bij sommige onderwerpen bepaalde gebaren niet thuishoren.

De boodschap van de auteurs over natuurlijkheid is soms ook dubbel. Mertens (1992) zegt weliswaar dat luisteraars natuurlijkheid verwachten, maar beveelt even verderop wel aan ‘als u aan het opsommen bent, en u duidt de opsomming mee met uw vingers aan, dan kan het publiek daar alleen maar baat bij vinden’ (p. 18). Vergelijkbaar geven Wiertzema en Jansen (2004) aan de ene kant gedetailleerde instructies over hoe te staan en te bewegen, maar stellen ze tegelijkertijd dat ‘ingestudeerde gebaren al snel overdreven en theatraal over komen’ (p. 119).

In een vergelijkbaar corpus van populaire Engelse adviesboeken is er meer overeenstemming. Het gebruik van gebaren wordt gekoppeld aan ‘to speak with power’ (Khan-Panni, 2009, p. 151). Atkinson (2004, p. 353) ziet het als een teken van ‘expressiveness, individuality and liveliness’ en hij betwijfelt dat het gebruik van gebaren zou afleiden van de inhoud van de presentatie. Het gebruik van betekenisvolle gebaren wordt de leerlingspreker voorghouden: een hakkende beweging maken met de zijkant van de hand om een verdeling uit te beelden of een-hand-over-het-hart als teken van een ‘deeply felt emotion’ (Leanne, 2009, p. 29). Oefenen wordt belangrijk gevonden: ‘Gestures have to be planned in advance so you can incorporate them during your speech rehearsal’ (Laskowski, 2001, p. 112).

Methode

Om een antwoord te vinden op de vraag of het wapperen van de handen vermeden moet worden (Van der Meiden, 1991) en een reden is om een training te volgen zoals weervrouw Willemijn Hoebert deed, en om het effect van gebaren te zien op retentie en sprekerswaardering, ontwikkelden we een experiment in een voor onze studenten aan de Technische Universiteit Delft gebruikelijke situatie: er werd een informatieve presentatie gegeven met PowerPoint-slides.

De presentatie en de gebaren

Als presentatie is gekozen voor een college van vijftien minuten (2107 woorden) over communicatie en gedragsbeïnvloeding.² Het college was informatief van opzet en paste bij de gebruikelijke onderwerpen bij ons communicatieonderwijs. Het college werd ondersteund met 23 PowerPoint-slides (inclusief animaties 32 klikmomenten).

De spreker werd gevonden in een ervaren universitair docent (man; 32 jaar). Hij presenteerde het college (met PowerPoint) voor de camera (Figuur 1) (de opnames zijn te zien via³). Er waren drie condities, waarin de volgende versies werden getoond.

Versie 01: zonder gebaren (Pilaar). De spreker behoudt gedurende de gehele presentatie dezelfde houding: de rustpositie (Kendon, 1980). De handen zijn voor de buik in elkaar gevouwen alsof ze een tennisbal vasthouden (Wiertzema & Jansen, 2004, p. 124) waardoor de spreker er 'als een pilaar' bijstaat.

Versie 02: alleen beat-gebaren (Beat). Tijdens de presentatie maakt de spreker beat-gebaren (gebaren die het ritme van gesproken woorden begeleiden en bekrachtigen). De instructie voor de spreker was dat hij vanuit de rustpositie (versie 01) regelmatig beide handen/onderarmen tegelijk/parallel op en neer zou bewegen; vervolgens moesten de handen/armen terugkeren in de rustpositie. De bewegingen tussen de rustposities telden als een complete beat (vergelijk Beattie & Coughlan, 1999, p. 46)

Versie 03: beat-gebaren afgewisseld met betekenisvolle gebaren (Mix). De spreker maakt naast beat-gebaren (versie 02), ook betekenisvolle gebaren (iconische, metaforische en deiktische gebaren). Hij maakt bijvoorbeeld (i) gebruik van gebaren die een concreet object/actie uitbeelden zoals een raam dichtdoen of kijken op een telefoon, (ii) gebaren die passen bij een abstracte tekst zoals 'anders', en (iii) gebaren die wijzen naar een plek op het scherm waar de PowerPoint-presentatie is te zien. Ook hier geldt dat alle bewegingen van de spreker tussen beide rustposities een compleet gebaar vormen.



Figuur 1 Weergave van spreker in rustpositie

De resulterende video's duren gemiddeld 14:56 minuut ($SD = 14$ seconden). De video's met gebaren bleken na controle gemiddeld 222 gebaren te bevatten. De beat-conditie bevatte 202 gebaren; de mix-conditie 243 gebaren (waarvan 128 beat-gebaren). Er was geen vierde conditie met alleen

betekenisvolle gebaren. Wanneer het aantal gebaren in de condities ongeveer gelijk zou worden gehouden, zou een spreker met circa 200 betekenisvolle gebaren in een video van 15 minuten gaan lijken op een mimespeler.

De gebaren lijken op het gebruikelijke presentatiegedrag van de spreker (vergelijk een openbaar optreden van de spreker in een uitzending van de *Universiteit van Nederland* (Wackers, 2016)). Het grootste verschil – door het experiment opgelegd – betreft de beperkte beweeglijkheid van de spreker (door een vaste positie op het scherm kon de spreker niet rondlopen). De intonatie en gezichtsuitdrukkingen lijken bij globale analyse per conditie niet systematisch verschillend.

Meetinstrumentarium (retentietest- en opvattingenvragenlijst)

We ontwikkelden een retentietest- en een opvattingenvragenlijst.⁴

Retentietest: een meerkeuzetest van dertig items (vier antwoordalternatieven per vraag) om de retentie te meten. In een eerder onderzoek waarbij hetzelfde materiaal werd gebruikt (Blokzijl & Andeweg, 2006) bleek de test een goede samenhang te hebben (Cronbach's $\alpha = .75$).

Opvattingenvragenlijst: een lijst met twaalf stellingen om de opvattingen over de spreker te meten. De stellingen gingen vergezeld van een vierpuntsschaal (helemaal niet mee eens; beetje mee eens; redelijk mee eens; heel erg mee eens).⁵ De stellingen hadden betrekking op drie factoren: de natuurlijkheid van de spreker, de boeiendheid van de presentatie, en de begrijpelijkheid. Zie Tabel 1.

Tabel 1 Factoren sprekerswaardering

<i>Natuurlijkheid (Cronbach's α .77)</i>
– De spreker weet de luisteraars te enthousiasmeren.
– De spreker komt ongedwongen over.
– De spreker presenteert levendig.
– Ik vind dat de spreker op natuurlijke wijze presenteert.
<i>Boeiendheid (Cronbach's α .80)</i>
– Ik vond het een boeiend college.
– Het was een interessant college.
– Ik kon mijn aandacht goed bij het college houden.
– Dit college maakte me nieuwsgierig naar meer informatie.
<i>Begrijpelijkheid (Cronbach's α .79)</i>
– De spreker benadrukt de belangrijkste zaken.
– Ik vond het een helder college.
– Ik vond het college inhoudelijk duidelijk.
– Dit college was begrijpelijk.

Proefpersonen

De drie varianten van het mini-college werden getoond aan groepen eerstejaars studenten Technische Bestuurskunde van de Technische Universiteit Delft ($N = 229$). Ze kregen een van de drie versies van het college gepresenteerd tijdens een practicum mondeling presenteren dat deel uitmaakt van het vak *Probleemanalyse*. De studentgroepen werden willekeurig over de drie condities verdeeld (Pilaar: $n = 71$; Beat: $n = 77$; Mix: $n = 81$), waarbij rekening gehouden werd met het moment van de dag (ochtend/middag) waarop het experiment plaatsvond en de betrokken begeleider (vijftien verschillende begeleiders, onder wie de spreker). Ter controle van de verdeling van de studenten over de drie condities werden drie stellingen voorgesteld (na de retentietest, vóór de opvattingenvragenlijst) met betrekking tot voorkennis, belangstelling voor het onderwerp en de noodzakelijkheid van de informatie in het college voor hun latere beroep. Een analyse van deze zelfrapportage liet zien dat de drie groepen niet significant van elkaar verschilden op deze aspecten ($F(6, 446) = 0.203$; $p = .976$). Uit een pre-analyse van de resultaten bleek dat de twee groepen die tijdens de afname van het experiment college hadden van de spreker significant (positief) afweken van de overige studentgroepen ten aanzien van de sprekerswaardering (ten aanzien van de retentietest waren er geen verschillen). Om vertekening van de resultaten te voorkomen, zijn de scores van deze groepen ($n = 17$) in de analyse van de sprekerswaardering niet meegenomen.

Resultaten

Het experiment leverde twee soorten testresultaten op.

Retentie – meerkeuzetest

De retentietest bestond uit dertig vragen. Het studentenresultaat werd teruggerekend naar een tentamencijfer tussen 1-10. Tabel 2 geeft een overzicht van de gemiddelden en de standaarddeviaties. In totaal scoorden de studenten gemiddeld 5,51 ($SD = 1,27$) (maximum: 8,67; minimum: 2,33).

Tabel 2 Retentie – resultaten meerkeuzetest

Conditie	Gemiddelde	<i>N</i>	<i>SD</i>
Pilaar	5,30	71	1,34
Beat	5,82	77	1,18
Mix	5,39	81	1,24
Totaal	5,51	229	1,27

Uit een univariate variantieanalyse blijkt een effect van conditie ($F(2, 226) = 3.765; p < .05$). Een Tukey-posttest laat zien dat de scores in de beat-conditie significant verschillen van die in de pilaar-conditie ($p = .032$). De verschillen tussen de mix-conditie en de overige twee condities zijn niet significant.

Sprekerswaardering – opvattingenvragenlijst

De studenten gaven hun mening – met behulp van een vierpunts-Likert-schaal – over drie sprekersaspecten (zie Tabel 1). De resultaten zijn terugerekend naar een score tussen 0-10, zie Tabel 3.⁶

Tabel 3 Sprekerswaardering

Conditie		Gemiddelde	SD	N
Natuurlijkheid	Pilaar	4,59 ^a	1,48	70
	Beat	5,70 ^b	1,74	67
	Mix	5,08	1,60	72
	Totaal	5,11	1,66	209
Boeiendheid	Pilaar	6,04	1,76	70
	Beat	6,52	1,62	67
	Mix	5,94	1,49	72
	Totaal	6,16	1,64	209
Begrijpelijkheid	Pilaar	7,66	1,38	70
	Beat	7,63	1,59	67
	Mix	7,99	1,24	72
	Totaal	7,77	1,41	209

* Verschillend superscript: $p < .01$

Uit een multivariate variantieanalyse blijkt een effect van conditie ($F(6, 408) = 4.340; p < .001$). Uit een nadere analyse blijkt dat dat effect moet worden toegeschreven aan de factor natuurlijkheid ($F(2, 206) = 8.206; p < .001$). Uit de Tukey-posttest blijkt dat het verschil in natuurlijkheid tussen de pilaar-conditie en de beat-conditie significant is ($p < .01$): de spreker in de beat-conditie scoort hoger dan zijn alter-ego in de pilaarconditie. De score voor natuurlijkheid in de mix-conditie verschilt niet significant van die in de beide andere condities. De verschillen tussen de condities bij de factor boeiendheid zijn niet significant ($F(2, 206) = 2.498; p = .085$). Hetzelfde geldt voor de verschillen bij de factor begrijpelijkheid ($F(2, 206) = 1.463; p = .234$).

Discussie en conclusie

In dit onderzoek is nagegaan in hoeverre het maken van gebaren tijdens een informatieve presentatie invloed heeft op de retentie (wat onthoudt

het publiek?) en de sprekerswaardering (hoe waardeert het publiek de spreker?). We onderscheidden drie condities: een spreker die geen gebaren maakt, een spreker die alleen beat-gebaren maakt en een spreker die naast beat-gebaren ook betekenisvolle (iconische, metaforische, deiktische) gebaren maakt (mix-conditie).

Voor de retentie van de informatieve presentatie blijkt dat luisteraars bij een spreker die zijn tekst begeleidt met beats significant meer informatie onthouden van de presentatie dan luisteraars bij dezelfde spreker die helemaal geen gebaren maakt. De mix-conditie neemt hier een tussenpositie in: het verschil met de beat-conditie of de pilaar-conditie is niet significant.

Wat betreft de waardering van de spreker blijkt dat de spreker in de beat-conditie significant ‘natuurlijker’ wordt gevonden dan die in de pilaar-conditie. Ook hier waren de verschillen tussen de mix-conditie enerzijds en de beat-conditie en de pilaar-conditie anderzijds niet significant. Aansluitend bij de bevindingen van Hostetter (2011), Kang et al. (2012) en Maricchiolo et al. (2009) was onze verwachting dat de mix-conditie met een spreker die gebaren maakt die concrete en abstracte objecten verduidelijkt en die wijsgebaren maakt afgewisseld met beat-gebaren, beter zou scoren dan de beat- en pilaar-conditie. Dat was dus niet het geval. De uitkomsten in ons experiment waren voor ons presentatiedocenten dan ook verwarrend. In onze cursussen adviseren we – in navolging van vele handboekschrijvers – het ondersteunen van betekenisvolle gebaren.

We hebben drie mogelijke verklaringen bij deze wat verwarrende uitkomst.

De gebaren in de mix-conditie waren ‘onnatuurlijk’. De spreker kreeg gebaren voorgeschreven om te maken (‘scripted gestures’). Hij moest bijvoorbeeld zijn vinger heen-en-weer bewegen bij een ontkennende stelling, bij een vraagstelling moest hij zijn beide handpalmen naar boven draaien. De presentatie werd enkele malen geoefend. Misschien was dit te weinig? Mogelijk werd deze versie daarom als minder spontaan (natuurlijk) bevonden dan de beat-conditie en werkten de gebaren daardoor niet als door ons verwacht. Het gedrag van de spreker lijkt op zijn gebruikelijke spreekgedrag (te observeren in Wackers, 2016). Uit ander onderzoek blijkt dat de effecten van ingestudeerdheid ook meevallen als we afgaan op Hostetter (2011), die in een meta-analyse van 38 studies liet zien dat de effecten van gebaren gelijk zijn bij spontane bewegingen en bewegingen die gescript zijn.

Er waren andere systematische verschillen tussen de condities. Mogelijk kan ook ander fysiek gedrag van de spreker van invloed zijn op de retentie van de luisteraars en hun waardering voor de spreker, bijvoorbeeld mimiek, intonatie, oogcontact, of bewegingen in de ruimte. Bij globale analyse leken

mimiek en intonatie echter niet systematisch verschillend. Verder is ons geen concreet onderzoek bekend waarin verschillen in expressie invloed zouden hebben op de retentie van een presentatie.

De gebaren in de condities verschilden in lengte. Hoewel er geen sprake was van significante verschillen tussen het *aantal* gebaren in de condities, viel wel op dat de gebaren in *tijdsduur* van elkaar verschilden: de ene keer werden er tussen twee rustposities meer bewegingen gemaakt dan de andere keer, oftewel, de tijdsduur van het gebaar verschilde. In andere onderzoeken verschilt bijvoorbeeld de manier waarop beats geteld zijn: Kang et al. (2012) zien een beat als één opwaartse beweging, terwijl in ander onderzoek de definitie van McNeill (1985, p. 359) wordt aangehouden, waarin het gaat om ‘kleine simpele bewegingen die snel worden uitgevoerd’. We besloten daarom de verschillen in tijdsduur van de gebaren nader te onderzoeken.

Met behulp van video-editingsoftware zijn de beide gebarenvideo's geanalyseerd om de timing en dosering van de gebaren beter in kaart te brengen. De software maakt het mogelijk om de video meerdere malen vertraagd (in dit geval op veertig procent van de oorspronkelijke snelheid) af te draaien: het werd daardoor mogelijk het begin en eind van de bewegingen tussen rustposities precies vast te leggen en daarmee de duur van het gebaar te bepalen. Tabel 4 geeft een overzicht van de verschillen in duur.

Tabel 4 Gebaarlengte verschillen

Conditie	Duur presentatie*	Totaal aantal gebaren	Gebaartijd	Gemiddelde duur gebaar*	SD
Beat	894	202	598	2,96#	1,58
Mix	917	243	528	2,17	1,45

* Duur in seconden; # verschil $p < .001$

Hoewel de spreker in de beat-conditie in absolute zin minder gebaren maakt dan zijn alter ego in de mix-conditie ($t(444) = 1.95$; $p = .052$), heeft hij zijn handen gemiddeld wel langer los (niet in rustpositie). Dat verschil tussen de beat-conditie en de mix-conditie is wel significant ($F(1, 443) = 32.143$; $p < .001$). Samenvattend blijkt dus dat de spreker in de beat-conditie 67 procent van de tijd zijn handen los heeft tegen 58 procent van de tijd in de mix-conditie (gebaartijd delen door duur presentatie), hoewel de spreker in de mix-conditie vaker de handen uit de rustpositie haalt. Een onbeantwoorde vraag is daardoor ook of de betere resultaten op retentie in de beat-conditie nu komen door het *type* gebaar (beat ten opzichte van betekenisvol) of door de *hoeveelheid tijd* die eraan besteed wordt (lang ten opzichte van kort). Vervolgonderzoek kan hier verder op inzoomen, bijvoorbeeld door

condities met dezelfde soort gebaren maar met variatie in gebarentijd te vergelijken.

Concluderend: in dit onderzoek onthielden luisteraars meer van wat een spreker vertelde bij het gebruik van beat-gebaren in vergelijking tot geen gebaren. Daarnaast werd de spreker als natuurlijker ervaren bij het gebruik van beat-gebaren. De waarschuwingen van auteurs over 'spinazie-hakkende' sprekers (Van der Meijden, 1991) die er met hun gedachten niet bij zijn (Krusche, 1986) lijken hiermee genuanceerd te worden. Het inschakelen van een presentatiecoach (zoals weervrouw Willemijn Hoebert deed) om wapperen te verminderen, is mogelijk niet de beste strategie om de presentatie van een informatieve voordracht te verbeteren.

Noten

We danken onze collega's van het Instituut voor Talen en Academische Vaardigheden (Technische Universiteit Delft), en met name Martijn Wackers, voor hun onmisbare inzet bij het experiment. Ook danken we de reviewers en de redactie voor hun stimulerende feedback.

1. In de laatste druk (2016; buiten corpus 1980-2010) wordt uitgebreider gesproken over gebaren.
2. Het college is gebaseerd op Pol, Swankhuisen en Fennis (2002). Zie Blokzijl en Andeweg (2006).
3. Op YouTube.com zijn de versies te bekijken: versie 01 (youtu.be/kal-YVNCnll); versie 02 (youtu.be/jOFvdjFBtng); versie 03 (youtu.be/YLXeFtQT1zA). Bij veel andere onderzoeken worden gebaren slechts beschreven of uitgelegd aan de hand van een foto. Het zou goed zijn als het gebruikte videomateriaal online staat.
4. De volledige retentietestvragenlijst en de opvattingenvragenlijst zijn op te vragen bij de auteurs.
5. De middenkeuze *geen mening* werd weggelaten. Hoewel een vierpunts-Likertschaal theoretisch is te motiveren (Garland, 1991), was de keuze praktisch omdat het standaardformulier dat digitaal is uit te lezen zo is opgezet.
6. Het afwijkende aantal proefpersonen wordt veroorzaakt door het weglaten van twee groepen (zie methode) en ontbrekende antwoorden bij het deel sprekerswaardering.

Bibliografie

- Atkinson, M. (2004). *Lend Me Your Ears*. London: Random House Ltd.
- Beattie, G., & Coughlan, J. (1999). An experimental investigation of the role of iconic gestures in lexical access using the tip-of-the-tongue phenomenon. *British Journal of Psychology*, 90, 35-56.
- Beattie, G., & Shovelton, H. (1999). Mapping the range of information contained in the iconic hand gestures that accompany spontaneous speech. *Journal of Language and Social Psychology*, 18, 438-462.

- Blokkzijl, W.J., & Andeweg, B.A. (2006). De effectiviteit van tekstdia's vergeleken met die van visualiserings in mondelinge presentaties. *Tijdschrift voor Taalbeheersing*, 28, 108-124.
- Blum, K. (1982). *Praktijkboek overtuigend spreken*. Amsterdam/Brussel: Intermediair.
- Boer, H. de (1986). *Doelmatige werkmethoden voor teksten schrijven, voordrachten houden, notulen maken*. Utrecht: Het Spectrum.
- Braas, C., Kat, J., Timmer, G., & Ville, I. (2001). *Presenteren*. Groningen: Wolters-Noordhoff.
- Eijk, I. van (1987). *De spreekhulp*. Amsterdam: Contact.
- Garland, R. (1991). The mid-point on a rating scale: Is it desirable. *Marketing bulletin*, 2, 66-70.
- Hostetter, A.B. (2011). When do gestures communicate? A meta-analysis. *Psychological bulletin*, 137, 297-315.
- Iverson, J.M., & Goldin-Meadow, S. (1998). Why do people gesture as they speak? *Nature*, 396, 228.
- Janssen, D. (Red.). (2002). *Zakelijke communicatie* (4^e geh. herz. dr.). Groningen: Wolters-Noordhoff.
- Kang, S., Hallman, G.L., Son, L.K., & Black, J.B. (2012). The different benefits from different gestures in understanding a concept. *Journal of Science Education and Technology*, 22, 825-837.
- Kendon, A. (1980). Gesticulation and speech: Two aspects of the process of utterance. In M. R. Key (Ed.), *The Relationship of Verbal and Nonverbal Communication* (pp. 207-227). The Hague: Mouton.
- Khan-Panni, P. (2009). *Stand and Deliver: Leave them Stirred not Shaken*. Academy Press: Penryn.
- Kirchner, B. (1983). *Spreken voor een groep. Techniek en praktijk nader uiteengezet*. Nijmegen: Dekker & Van de Vegt.
- Krauss, R.M., Morrel-Samuels, P., & Colasante, C. (1991). Do conversational hand gestures communicate? *Journal of Personality and Social Psychology*, 61, 743-754.
- Krauss, R.M., & Hadar, U. (1999). The role of speech-related arm/hand gestures in word retrieval. In R. Campbell & L. Messing (Eds.), *Gesture, Speech, and Sign* (pp. 93-116). Oxford: Oxford University Press.
- Krusche, H. (1986). *Neem het woord*. Baarn: Bigot & Van Rossum.
- Laskowski, L. (2001). *10 Days to More Confident Public Speaking*. New York, NY: Warner Books.
- Leanne, S. (2009). *Say it like Obama. The Power of Speaking with Purpose and Vision*. New York, NY: McGraw-Hill.
- Maricchiolo, F., Gnisci, A., Bonaiuto, M., & Ficca, G. (2009). Effects of different types of hand gestures in persuasive speech on receivers' evaluations. *Language and Cognitive Processes*, 24, 239-266.
- McNeill, D. (1985). So you think gestures are nonverbal? *Psychological Review*, 92, 350-371.
- McNeill, D. (1992). *Hand and Mind*. Chicago, IL: The University of Chicago Press.
- Meiden, A. van der (1991). *Over spreken gesproken*. Groningen: Wolters-Noordhoff.
- Mertens, V. (1992). *Spreken voor publiek*. Leuven/Apeldoorn: Garant.
- Pol, L.R., Swankhuisen, C.E., & Fennis, B.M. (2002). *Communicatie en gedragsbeïnvloeding*. Alphen aan den Rijn: Kluwer.
- Ritter, P.H. jr. (1940). *Van stamelaar tot redenaar*. Utrecht: Erven J. Bijleveld.
- Sanders, P. (1987). Enthusiasm Awareness in Experiential Learning. *Organizational Behavior Teaching Review*, 11(4), 21-30.
- Sueyoshi, A., & Hardison, D.M. (2005). The role of gestures and facial cues in second language listening comprehension. *Language Learning*, 55, 661-699.
- Steehouder, M., Jansen, C., Maat, K., Staak, J. van der, Vet, D. de, Witteveen, M., & Woudstra, E. (2006). *Leren communiceren* (5^e herz. dr., bewerkt door M. Gijzen). Groningen: Wolters-Noordhoff.
- Waarlo, N. (2017, 2 december). Het perfecte ... weerbericht. *De Volkskrant*, p. 5.
- Wackers, M., Jong, J.C. de, & Andeweg, B. (2016). Structureren om onthouden te worden: Retentietechnieken in presentaties van onderzoekers en toespraken van politici.

- In D. van de Mierop, L. Buysse, R. Coesemans, & P. Gillaerts (Red.), *De macht van de taal: Taalbeheersingsonderzoek in Nederland en Vlaanderen* (pp. 295-308). Leuven: Uitgeverij Acco.
- Wackers, M. (2016). *Leren leren: Hoe vertel jij een onvergetelijk verhaal?* Universiteit van Nederland. Geraadpleegd op 1 januari 2019, youtu.be/ODQdvWN67c8
- Wagner, P., Malisz, Z., & Kopp, S. (2014). Gesture and speech in interaction: An overview. *Speech Communication*, 57, 209-232.
- Wiertzema, K., & Jansen, P. (2004). *Spreken in het openbaar* (2^e dr.). Amsterdam: Pearson Education.

Opvattingen van mbo-docenten over beroepsgerichte taal- en communicatie-eisen en de aansluiting op een dynamische arbeidsmarkt

Annet Bruin¹, Jacqueline van Kruiningen¹, Kees de Glopper¹, Arjan van der Meijden² en Nicole van der Meulen¹

¹ Rijksuniversiteit Groningen

² Kohnstamm Instituut

TVT 41 (1): 27–40

DOI: 10.5117/TVT2019.1.002.BRUI

Abstract

Tailoring to job demands: Language arts in vocational education

This article explores two cases of teacher beliefs about the connection between students' language and communication skills in secondary vocational education and training (in the Netherlands: mbo) and the skills required in the dynamic labour market. Previous studies showed that alumni's language and communication skills do not fit in with changing demands of the labour market. This study pertains to language and subject teachers from two vocational curricula who discussed this topic in focus group meetings. These data were complemented with one-to-one interviews and a survey among language teachers. The findings show that the contacts between the programs and employers are substantial, but that teachers are in need of more information about future demands related to communication skills in changing and new professions. Also, it appears to be difficult for language teachers to design a language integrated curriculum in cooperation with subject teachers. Finally, the national requirements for the examination of language skills collide with the teachers' and programs' desire to design and assess authentic professional tasks. These results lead to insights that feed the development of knowledge about language teaching in vocational training and about ways to connect vocational curricula to changing demands of the (future) labour market.

Keywords: Secondary vocational education and training (VET), language and communication skills, teacher beliefs, labour market

Inleiding

Het mbo bereidt leerlingen voor op een beroep. Dit is niet zo eenduidig als het lijkt: het beroepenveld is aan veel veranderingen onderhevig en dat stelt specifieke eisen aan de afgestudeerde. Vakspecificiteit is een voorwaarde om een opleiding goed aan te laten sluiten op de beroepspraktijk, maar ook taal- en communicatievaardigheden zijn van essentieel belang voor het succes van een mbo'er op de arbeidsmarkt. Deze vormen steeds vaker onderdeel van die vakspecificiteit. Professionele beroepsuitoefening krijgt in belangrijke mate gestalte in en door taalgebruik (Beaufort, 2009). Voorbereiding daarop in het onderwijs is daarom van groot belang.

Praktijkervaringen en onderzoek binnen het mbo maken echter duidelijk dat er aan *teaching for transfer* van taal- en communicatiecompetenties nog vele vragen kleven. Binnen het mbo speelt de problematiek van aansluiting op de eisen van de veranderende arbeidsmarkt al geruime tijd. Studies in het eerste decennium van de 21^{ste} eeuw toonden aan dat de taalbeheersing van mbo-studenten onvoldoende zou zijn om de opleiding succesvol af te ronden en te functioneren in het beroep (o.a. Neuvel, Bersee, Den Exter, & Tijssen, 2004; Onderwijsraad, 2006). Ondanks het feit dat naar aanleiding van deze studies in 2007 taaleisen in de kwalificatiedossiers voor de verschillende opleidingen werden vastgesteld, het vak Nederlands – na een periode van afwezigheid – geleidelijk aan weer terug is gekomen in het mbo, en er verplichte examens zijn ingesteld, toont onderzoek nog steeds geen rooskleurig beeld van het taalvaardigheidsniveau van mbo-studenten (Raaphorst & Steehouder, 2010). Ook geven werkgevers aan dat ze ontevreden zijn over de communicatieve vaardigheden van studenten (Hendrix, Hovens, & Kappers, 2012; Van Knippenberg, 2012).

Daarbij is het voor mbo-instellingen lastig om de toekomstgerichtheid van het opleidingsaanbod op orde te houden, zo blijkt uit actueel onderzoek van het Sociaal en Cultureel Planbureau; niet alleen vanwege grote veranderingen op de arbeidsmarkt, maar ook omdat voor de grote variëteit aan opleidingsrichtingen en 'soorten vakmanschap' een even zo grote variëteit aan toekomstperspectieven bestaat (Buisman & Van der Velden, 2017; Meng, Van der Meijden, Levels, Van Schooten, & Van der Velden, 2017; Turkenburg & Vogels, 2017). Volgens het rapport van het SCP werken mbo-opleidingen nog te weinig vraaggericht en zijn ze te weinig wendbaar.

Opleidingsteams hebben vaak geen scherp beeld van onder meer de taal- en communicatiecompetenties die van belang zijn voor de inzetbaarheid op korte en op langere termijn.

In een consortium van vier mbo-scholen en twee onderzoeksinstituten wordt nu onderzocht welke taal- en communicatiecompetenties nodig zijn op de dynamische arbeidsmarkt en welke opleidingspraktijken bijdragen aan het vermogen van opleidingsteams om in te spelen op de eisen die de dynamische arbeidsmarkt stelt. Dit door het NRO gesubsidieerde project krijgt vorm in een reeks casestudies waarin bovenstaande problematiek wordt onderzocht, en op basis waarvan opleidingen werken aan (her-) ontwerp van hun onderwijs. In de startfase van het project vonden focus-groepgesprekken en interviews plaats met taal- en vakdocenten. Ook is een enquête gehouden onder docenten Nederlands. In dit artikel rapporteren we op basis van twee masterscripties (Bruin, 2017; Van der Meulen, 2017) over de uitkomsten van deze verkenning bij twee opleidingsteams.

Ontwikkelingen in onderwijs en beroepenveld

De huidige situatie op de arbeidsmarkt laat zien dat banen onder invloed van digitalisering, flexibilisering, globalisering en technologische ontwikkelingen in snel tempo veranderen van inhoud en aard, of zelfs verdwijnen (CPb, 2015; Christoffels & Baay, 2016; Turkenburg & Vogels, 2017). Vanzelfsprekend hebben deze veranderingen ook gevolgen voor de vereiste taal- en communicatiecompetenties van werknemers. Uit onderzoek van Turkenburg en Vogels onder mbo-professionals blijkt dat zij communicatieve vaardigheden het meest van belang vinden voor de toekomst van de studenten (met name in de sectoren zorg, welzijn en economie), naast vakspecifieke en sociaal-emotionele vaardigheden, probleemoplossend vermogen, kritisch denken en digitale geletterdheid, en basisvaardigheden als taalvaardigheid en burgerschapsvaardigheden (Turkenburg & Vogels, 2017). Goede taal- en communicatievaardigheden zijn in dit palet dus van groot belang. Een taal- en communicatiecompetente student heeft meer perspectief op doorstroming naar een hoger onderwijsniveau en is beter voorbereid op de kennissamenleving (Buisman & Van der Velden, 2017; Meng et al., 2017; Turkenburg & Vogels, 2017). Uit onderzoek van de Samenwerkingsorganisatie Beroepsonderwijs Bedrijfsleven (SBB) blijkt dat veel leerbedrijven ontevreden zijn over de taal- en communicatievaardigheden van de huidige mbo-studenten (Onderwijs in cijfers, 2016).

Goede voorbereiding op de beroepsuitoefening behelst enerzijds een gerichtheid op generieke communicatievaardigheden, en anderzijds op vaardigheden voor gespecialiseerd taalgebruik (Elbers, 2012). Turkenburg

en Vogels (2017) benadrukken, in navolging van de OECD, het belang van *transferable skills*; vaardigheden die ‘overdraagbaar’ zijn tussen beroepen. Tegelijkertijd brengen verschillende beroepen uiteenlopende taaleisen met zich mee. Taalcompetenties maken steeds vaker onlosmakelijk deel uit van vakmanschap. Beroepen kennen hun eigen, vakspecifieke taal of register en hun vakspecifieke professionele manieren waarop taal – mondeling en schriftelijk – wordt gebruikt en vakkennis wordt verwoord (Bawarshi & Reiff, 2010; Biber & Conrad, 2009; Elbers, 2012; Faber-de Lange & Van der Pool, 2015).

Taal- en communicatiecompetenties zijn niet alleen van belang voor beroepsuitoefening, maar ook voor leren. Tussen taalgebruik, denken en leren bestaan immers intrinsieke verbindingen (Brandt, 2009; Britton, 1970; Hajer & Meestringa, 2004; Klein, 1999; Mercer, 1995). Docenten gebruiken taal als middel om vakinhoud over te brengen. Studenten moeten ‘de taal van het vak’ (mondeling en schriftelijk) leren beheersen om de vakinhoud te leren en te verwerken. Taal leren is dus niet alleen functioneel voor taalproductie en communicatie; taal dient ook als leer- en denkinstrument. Waar taal, communicatie en vakkennis hand in hand gaan, wordt de leeropbrengst vergroot (Bolle & Van Meelis, 2014; Elbers, 2012; Van Knippenberg, 2010).

Dit uitgangspunt komt in het onderwijs tot uitdrukking in vormen van taalgericht vakonderwijs. Uit onderzoek blijkt dat opleidingsprogramma's waarin taal- en vakonderwijs op elkaar afgestemd worden, succesvol zijn en voor betere resultaten zorgen dan wanneer het vak Nederlands losstaat van de beroepsvorming (Elbers, 2012). Meng et al. (2017) trokken vergelijkbare conclusies met betrekking tot het al dan niet verbinden van onderwijs in vakspecifieke competenties met onderwijs in generieke vaardigheden. Uit verschillende studies blijkt echter dat aandacht voor taal in de vaklessen in het mbo tekortschiet (Bolle, 2011; Elbers, 2012; Hajer, 2015; Raaphorst & Steehouder, 2010; Van Knippenberg, 2010). Slag 1 (aandacht voor taal in alle vakken) van de *Drieslag Taal* (Bolle, 2009) – voor veel scholen uitgangspunt van hun taalbeleid – komt niet goed van de grond. Docenten hebben verschillende opvattingen over de rol van taal in hun onderwijs. Dat begint met de betekenis van het concept taalvaardigheid. De beïnvloedbaarheid van de ontwikkeling van taal- en communicatievaardigheden is een tweede punt (Hajer, 2005; Van Knippenberg, 2010). Naast de veelal lage verwachtingen die mbo-docenten hebben over de taalprestaties van studenten, zijn mbo-docenten veelal van mening dat zij studenten hoofdzakelijk opleiden voor vakmanschap. Een goede taalbeheersing vinden zij minder belangrijk (Van Knippenberg, 2010). Inzicht in de rol van actief taalgebruik bij leren ontstaat kennelijk niet vanzelf in de praktijk van ervaren docenten, zo stelt Hajer (2005).

Ook onder mbo-studenten wordt het belang van taal- en communicatiecompetenties nauwelijks onderkend; ze lijken weinig interesse te hebben voor het vak Nederlands. Uit een verkenning van het Sociaal en Cultureel Planbureau blijkt dat studenten vooral intrinsiek gemotiveerd raken door de essentie van het vakmanschap (Turkenburg & Vogels, 2017). Minder motiverend zijn de algemeen vormende vakken zoals Nederlands, die ver van de beroepsvorming af staan.

Volgens Van Knippenberg (2010) zullen mbo-docenten zich meer bewust moeten worden van hun rol bij de ontwikkeling van taal- en communicatievaardigheden van studenten. Een vereiste voor een adequate voorbereiding op professionele beroepsuitoefening is verder dat opleidingen een actueel beeld krijgen van de taal- en communicatiecompetenties die de (toekomstige) beroepen vereisen. Juist op dit gebied ontbreekt het aan kennis.

Het project Taalcompetenties voor een dynamische arbeidsmarkt

In het onderzoeks- en onderwijsontwikkelproject *Taalcompetenties voor een dynamische arbeidsmarkt* werken opleidingsteams van vier mbo-instellingen en twee onderzoeksinstituten in een reeks casestudies aan geïnformeerd (her)ontwerp van onderwijs (looptijd: september 2017-september 2020). De participerende opleidingsteams vinden het belangrijk dat hun studenten voldoen aan de taaleisen die de arbeidsmarkt stelt en ervaren de behoefte om hun responsiviteit te vergroten. Vanuit dat perspectief wordt gekeken naar ontwikkelingen in het beroep, wat die voor consequenties hebben voor de vereiste taalcompetenties van mbo-professionals, om vervolgens een vertaalslag te maken naar het onderwijs.

In de eerste projectfase wordt onderzocht welke visies mbo-docenten hebben op de taal- en communicatie-eisen die de arbeidsmarkt stelt, welke opvattingen zij hebben over de feitelijk en wenselijke aansluiting van het curriculum op die eisen, en wat zij doen in het beroepsonderwijs om aan die eisen tegemoet te komen. Een tweede, nog te starten onderzoekscomponent bestaat uit interviews met werkgevers. In dit artikel richten we ons op basis van twee casestudies op de visies van taal- en vakdocenten in twee opleidingen. In het kader van scriptieonderzoek van Bruin (2018) en Van der Meulen (2018) rapporteren we over de uitkomsten van twee focusgroepgesprekken met taal- en vakdocenten uit de opleidingsteams van de opleidingen Handel en Ondernemerschap (ROC Friese Poort, Leeuwarden) en van de opleiding Chauffeur Wegvervoer (Friesland College, Leeuwarden), en een breder uitgezette enquête onder docenten Nederlands.

Aan het focusgroepgesprek bij de opleiding Handel en Ondernemerschap (H&O) namen zes docenten deel: drie docenten Nederlands, een docent Engels en twee vakdocenten. Aan het focusgroepgesprek in de Chauffeursopleiding namen vier docenten deel: twee docenten Nederlands en twee vakdocenten. Omdat in het focusgroepgesprek met de docenten Handel en Ondernemerschap, dat tevens een project-introducerend karakter had, door tijdgebrek niet alle thema's even diepgaand konden worden besproken, zijn de data uit dat gesprek naderhand aangevuld met schriftelijke vragen onder de zes gespreksdeelnemers plus aanvullend nog drie andere vakdocenten van dezelfde opleiding. In het focusgroepgesprek bij de Chauffeursopleiding konden wel alle thema's even uitputtend worden besproken. Beide focusgroepgesprekken duurden twee uur.

Tijdens de focusgroepgesprekken had één onderzoeker de rol van gespreksleider; twee onderzoekers waren als observator aanwezig en maakten aantekeningen. Tevens zijn audio-opnames van de gesprekken gemaakt. De gespreksleider hanteerde een gespreksleidraad met dertien deelonderwerpen, die onderverdeeld waren in drie hoofdthema's: 1) de kennis over en visies van de gespreksdeelnemers over taal- en communicatie-eisen op de arbeidsmarkt, 2) de manieren waarop de docenten in hun onderwijs proberen aan die taaleisen tegemoet te komen, en 3) hun opvattingen over de aansluiting van hun onderwijsprogramma op de taaleisen van de arbeidsmarkt (en mogelijkheden/wensen om hieraan te werken). Deelthema's waren onder andere: visies op en ervaringen met taalgericht vakonderwijs en visies op de rollen en verantwoordelijkheden van taal- en vakdocenten in het taal- en communicatieonderwijs.

Voor de scriptie van Van der Meulen, die specifiek gericht was op taalexaminering in de mbo-opleidingen, bestond de dataverzameling uit dezelfde twee focusgroepgesprekken plus een schriftelijke enquête onder docenten Nederlands van beide onderwijsinstellingen over de taalexamenpraktijken in de opleidingen. Deze enquête is uitgezet onder 30 docenten Nederlands van ROC Friese Poort en 40 docenten Nederlands van het Friesland College. Van ROC Friese Poort hebben veertien docenten de enquête ingevuld; van het Friesland College dertien. In de enquête is, voornamelijk in de vorm van open vragen, gevraagd naar de ervaringen en opvattingen van de docenten over 1) de examinering van Nederlandse taalvaardigheid, 2) de normen voor taalvaardigheid in het referentiekader Nederlandse taal, en 3) de aansluiting van de taalexaminering op de arbeidsmarkt.

Alle data uit deze verkennende, grotendeels kwalitatieve studies zijn gecodeerd en geanalyseerd met behulp van het softwareprogramma *ATLAS.ti*. Voor de analyses in beide studies is een beoordelaarscontrole uitgevoerd

op een selectie van de data. De interbeoordelaarsbetrouwbaarheid van de analyses in de scriptie van Bruin was 79%; die van de analyses in de scriptie van Van der Meulen 93 %.

Eerste bevindingen op basis van twee casestudies

Beroepsgerichte taaltaken

Uit de focusgroepgesprekken wordt duidelijk dat de docenten van beide opleidingen veel contact hebben met het beroepenveld. Zowel de taal- als vakdocenten van beide opleidingen voeren stagegesprekken met leerbedrijven, bezoeken open dagen, organiseren projecten; werkgevers verzorgen gastcolleges en hebben zitting in onderwijsadviesraden. Vanuit de opleiding Handel en Ondernemerschap is er ook structureel contact met de ondernemers die de bevoorrading van outletwinkels van de school verzorgen.

Zowel de taal- als vakdocenten geven aan dat ze vinden dat ze een redelijk beeld hebben van ontwikkelingen in het vakgebied. Ook zien ze het belang van taal- en communicatieonderwijs in het licht van deze ontwikkelingen: de wereld wordt steeds taliger en samenwerking en transparante communicatie steeds belangrijker. Ze vinden dat ze behoorlijk zicht hebben op beroepsgerelateerde taal- en communicatie-eisen, maar uiten ook de behoefte om beter te weten wat die eisen zijn in de veranderende beroepen, en hoe ze hun onderwijs daarop kunnen inrichten. Bijvoorbeeld: worden er nog zakelijke brieven geschreven; moet je als werknemer een blog en een facebookbericht kunnen schrijven, en hoe richt je je onderwijs daarop?

Met name de digitalisering heeft gevolgen voor taal en communicatievereisten. Voor het vakgebied Handel en Ondernemerschap betreft dat bijvoorbeeld het gebruik van *social media* als marketingstrategie, met als gevolg dat klanten beter voorbereid naar de winkel komen. Dit betekent dat het onderwijs meer gericht moet zijn op professionele communicatie via *social media* en op een meer adviserende rol van verkopers. Zowel de taal- als vakdocenten merken dat studenten hier tegen problemen aanlopen, bijvoorbeeld rondom de vraag hoe formeel/informeel ze zich kunnen of moeten opstellen in professionele communicatie via *social media*; ze zijn immers informele communicatie via *social media* gewend. Ook de komst van webshops stelt nieuwe eisen aan verkoopspecialisten. Docenten geven aan dat toekomstige verkopers het jargon van hun vakgebied moeten beheersen en over de grenzen van hun vak met klanten moeten kunnen communiceren. Ze moeten leren klanten aan zich te binden door een verbaal slag te maken van technisch vakjargon naar voordelen voor de klant.

Een vakdocent Handel en Ondernemerschap geeft een voorbeeld van een verkoper in een bouwmarkt die vroeger meer een ‘veredelde vakkenvuller’ was die met goederen bezig was en weinig met verkoopgesprekken, maar door de transformatie van bouwmarkten naar woonwinkels nu ook een communicatief vaardige verkoper moet worden.

Bij het chauffeursberoep zorgt de ontwikkeling van onder andere interactief rijden voor een afname in werkgelegenheid. Studenten moeten zich meer dan voorheen kunnen presenteren en onderscheiden van concurrenten. Een docent Nederlands van de chauffeursopleiding zou bijvoorbeeld veel meer aan spreken en presenteren willen doen, waaronder: een professioneel, zakelijk voicemailbericht inspreken. Nu hoort hij soms ‘dingen dat je denkt nou ja, als dat nou het begin is van je visitiekaartje; dat kan wel anders’. Een vakdocent stelt: ‘het is voor die jongens of meiden van essentieel belang dat ze zichzelf ook leren te verkopen, dat ze hun eigen visitekaartje zijn en dat van het bedrijf, en trots zijn op datgene wat ze doen’. Ook benadrukken de docenten het belang van het visitekaartje voor het bedrijf, en in dat kader het belang van het eerste klantcontact. Een van de twee docenten Nederlands: ‘Nou, dan moet je je kunnen gedragen, zodat die klant denkt van “nou, leuke vent, goed bedrijf, ik word gehoord en serieus genomen”’. Studenten weten vaak niet hoe ze over (zouden moeten) komen en wat gepast is. Het taal- en communicatie-onderwijs zou gericht moeten zijn op bewustmaking van het belang van dit soort basale vaardigheden, concluderen de gespreksdeelnemers van de chauffeursopleiding.

Taalonderwijs en taalgericht vakonderwijs

Voor beide opleidingen geldt dat zowel vakdocenten als docenten Nederlands vinden dat hun huidige taalonderwijs onvoldoende aansluit op bovengenoemde ontwikkelingen, en dat het taalniveau van de studenten onvoldoende aansluit op de taaltaken die de arbeidsmarkt en overigens ook het hbo vereisen. Vakdocenten geven aan onvoldoende zicht te hebben op het taalniveau van de studenten.

Bij beide opleidingen is er behoefte aan intensievere en meer beroepsgerichte training van schrijf- en gespreksvaardigheden, waarbij vooral mondelinge klantcontacten worden gezien als centrale taaltaken. Ook leesvaardigheid wordt genoemd als een essentiële vaardigheid die vaak nog onvoldoende wordt beheerst door de studenten. De docenten zien dat veranderingen in de beroepen, waarbij steeds meer handwerk verdwijnt, leiden tot hogere eisen op dit terrein. Voor de chauffeursopleiding bijvoorbeeld: het kunnen lezen vrachtbrieven en andere documentatie op een boardcomputer.

Overigens wordt het belang van receptieve vaardigheden in beide gesprekken ook genoemd in het kader van ‘mee kunnen doen in de maatschappij’: goed burgerschap en de persoonlijke ontwikkeling van de studenten, ongeacht het beroep waarvoor wordt opgeleid. (Vergelijk de drieledige doelstelling van het mbo: beroepskwalificatie, doorstroom in de onderwijskolom en voorbereiding op burgerschap.) Daarbij wordt opgemerkt dat een gebrekkige leesvaardigheid regelmatig tot problemen bij de examinering leidt, met name bij de examens van de zaakvakken. Ook merken de docenten op dat studenten het belang van leesvaardigheid veel minder inzien. Toch worden de productieve taaltaken (mondelinge en schriftelijke vaardigheden) bij beide opleidingen, en vooral professionele gespreksvoering, gezien als meest relevante taaltaken.

De docenten Nederlands van beide opleidingen geven in de focusgroep-gesprekken aan dat ze waar mogelijk beroepsgerichte taaltaken integreren in het vak Nederlands, geregeld ook in samenwerking met vakdocenten. De vakdocenten in de focusgroepgesprekken tonen zich ook betrokken maar de verantwoordelijkheid voor het taalonderwijs leggen ze toch vooral bij de docent Nederlands. Opvallend is wel dat zowel taal- als vakdocenten van beide opleidingen in de focusgroepgesprekken aangeven dat ze vinden dat vakdocenten een grotere rol zou moeten spelen bij de taalvaardigheidsontwikkeling van studenten, en er een duidelijkere koppeling tot stand moet komen tussen het vak Nederlands en beroepsvorming, maar dat uit de enquête onder de bredere groep docenten Nederlands een heel ander beeld naar voren komt. Hun ervaring is dat vakdocenten het vak Nederlands en de eisen die op het terrein van taal- en communicatievaardigheden worden gesteld, niet belangrijk vinden: ‘Nee geen belangstelling voor’; ‘Nederlands wordt in mijn situatie niet breed gedragen’; ‘ze hebben geen idee bij ons’; ‘zij zijn bezig met hun eigen vakontwikkelingen, het heeft voor hen geen prioriteit’; ‘taal is nog niet van ons allemaal’. Slag 1 (aandacht voor taal in alle vakken) van de *Drieslag Taal* (Bolle, 2009) komt bij beide onderwijsinstellingen, en ook bij de twee opleidingen uit de casestudies niet goed van de grond.

Een ander aandachtspunt dat in de enquête onder de docenten Nederlands naar voren komt is dat ze vinden dat veel vakdocenten zelf het vereiste taalniveau niet beheersen, en dus ook het taalgebruik van hun studenten niet goed kunnen beoordelen. Het gaat daarbij vooral om indrukken: ‘er wordt hierover te weinig gesproken’, en: ‘onduidelijk is hoe er wordt gedacht en gehandeld’. Taal- en vakdocenten zijn dus niet heel bekend met elkaars onderwijspraktijken.

Ondanks de gesignaleerde problemen vinden de meeste docenten Nederlands in de enquête wel dat zowel taal- als vakdocenten een belangrijke rol moeten spelen in het taalonderwijs: 'Taal is de rode draad door elke opleiding en dient brede aandacht te krijgen bij alle lessen'. Immers: 'vakdocenten luisteren naar het taalgebruik van de student [...] zoals klanten, collega's en leidinggevenden ernaar luisteren'. De enquête-respondenten benadrukken dat door een koppeling van het taal- aan het vakonderwijs de motivatie van studenten en hun inzicht in het belang van taal kan worden vergroot. Een respondent: 'de student moet het ergens aan kunnen koppelen en hij moet er het voordeel van inzien. Ik denk dat je daaraan moet werken'.

In de focusgroepgesprekken wordt geopperd om het werkveld (nog) meer te betrekken bij invulling van het onderwijs en bij (het ontwerp van) de examinering. Bijvoorbeeld door een structurele oefening en toetsing van taaltaken in de beroepspraktijkvorming (stage), gastlessen van professionals over taal en communicatie in het beroep, en een structurele peiling van behoeften van werkgevers ten aanzien van taaleisen. Toch zijn er ook twijfels bij het betrekken van professionals uit het werkveld bij het onderwijs. Enerzijds omdat vakdocenten in de opleiding al voldoende input vanuit het werkveld kunnen leveren, en anderzijds omdat, zoals een docent Nederlands in het focusgroepgesprek Handel en Ondernemerschap aangeeft, ook wel wordt getwijfeld over het taalniveau van professionals.

Examinering

De examinering van het vak Nederlands is een veelbesproken onderwerp in de mbo-wereld. Deze examinering is opgesplitst in een centraal landelijk examen voor de receptieve vaardigheden en instellingsexamens voor de productieve vaardigheden. De centrale examens, met toetsing op basis van lees, luister- en kijk-luistervragen aan de hand van zakelijke teksten, zijn inhoudelijk gelinkt aan beroepsoverstijgende thema's rondom het vak Loopbaan & Burgerschap. Een voordeel van deze generieke examinering is dat de centrale aanpak de docenten ontlast, en dat deze per opleidingsniveau gelijk is voor alle studenten. Een veel gehoord kritiekpunt is dat de onderwerpen niet aanspreken en niet aansluiten bij de beleavingswereld van de studenten. Veel docenten, zowel in de focusgroepgesprekken als in de enquête, pleiten voor meer aansluiting bij de vakgebieden; een enkeling (in de enquête) pleit juist voor de generieke aanpak, omdat 'het referentiekader van een leerling vergroot mag worden', en omdat het werkveld breed is.

De verantwoordelijkheid voor examinering van de productieve taalvaardigheden (schrijven, spreken, gesprekken voeren) is neergelegd bij de scholen. Door de instellingsexamens zelf te maken, weliswaar met in achtneming van landelijke kwaliteitseisen, kunnen deze examens meer beroepsgericht worden ingevuld en daardoor in een voor de student relevante context worden geplaatst. Toch worden veel instellingsexamens generiek ingekocht om zo de beheerslast te verkleinen (Speijers, 2012). Om die reden wordt door docenten van ROC Friese Poort de beroepsgerichtheid in de examens gemist. Er is geen mogelijkheid om de studenten te beoordelen in authentieke of gesimuleerde beroepssituaties.

Op het Friesland College is men verder met het invoeren van zelf ontwikkelde, meer beroepsgerichte examinering. De ervaringen zijn overwegend positief. Uit de enquête: 'de onderwerpen zijn herkenbaarder'. Daar tegenover staan opvattingen van docenten die toch een meer centrale aanpak wensen omdat de inhoud per opleiding sterk kan verschillen en sommige scholen ermee 'aan de haal gaan'. Een terugkomende kritiek op de zelf te ontwikkelen instellingsexamens is vooral dat ze lastig te organiseren en te ontwerpen zijn. Er wordt veel van docenten gevraagd; de organisatie moet beter: 'je vergt veel van docenten als elke mbo-vakrichting een eigen IE moet ontwerpen'. En: 'de examendruk ontnemt me veel creativiteit om geïntegreerd in te kunnen spelen op beroepsgerelateerde taalproducten'. De controleprocedures binnen de mbo-instellingen ontnemen de docenten veel vrijheid.

Niet alleen het ontwerpen van de instellingsexamens wordt door docenten Nederlands als moeilijk ervaren, ook de beoordeling wordt als lastig ervaren; vooral het onderdeel gespreksvoering. Het ontbreekt de opleidingen bij beide onderwijsinstellingen aan capaciteit om met meer beoordelaars en in meer authentieke beroepssituaties te toetsen. Ook bij voorbereiding op het examen is er een capaciteitsprobleem: te veel studenten per groep, weinig lesuren, geen ruimte om één-op-één te oefenen. De docent Nederlands van de opleiding Handel & Ondernemerschap signaleert in het focusgroepgesprek dat de cijfers voor het examen nu vaak afhankelijk zijn van wat studenten toevallig tijdens heel verschillende stages leren. Zo leren studenten in een kledingwinkel beter een verkoopgesprek te voeren dan studenten in een supermarkt. Hij benadrukt het belang van leren op de werkplek: 'ze kunnen beter oefenen of dat soort dingen beter leren in een winkel dan bij ons'.

Conclusies en discussie

Deze kleinschalige studie biedt zicht op de praktijken en visies van taal- en vakdocenten in twee mbo-opleidingen aan twee mbo-scholen. Op basis van twee exploratieve casestudies is een beeld geschetst van de visies van mbo-docenten op taaltaken in de opleiding en in het beroepenveld, en geven hun opvattingen weer over de feitelijke en wenselijke aansluiting van het taalonderwijs op de arbeidsmarkt. Dit beeld bevestigt bevindingen uit eerder onderzoek over taalgericht vakonderwijs in het mbo grotendeels. Het blijkt lastig om taal- of taalrijk onderwijs en taaltoetsing met beroepsgereleerde taaltaken, en in samenwerking met vakdocenten vorm te geven. In tegenstelling tot wat blijkt uit eerdere studies, geven de docenten verder aan dat er vanuit de opleiding veel contact is met het werkveld en dat er ook wel zicht is op de taaleisen die de arbeidsmarkt stelt. Tegelijkertijd is er behoefte aan meer kennis over taaleisen in het veranderende werkveld. Wat in deze studie nadrukkelijk naar voren komt is dat de inhoud en opzet van de centraal georganiseerde examens (receptieve vaardigheden) en ook die van de instellingsexamens (productieve vaardigheden) vaak botsen met wat de opleidingen eigenlijk zouden willen, namelijk meer praktijkgerichte en minder generieke toetsing.

Om een vollediger zicht te kunnen bieden op de aard van deze aansluitingsproblematiek, en tot aanbevelingen te kunnen komen voor taalbeleid in het mbo, is aanvullend onderzoek nodig. In het project dat nu gestart is bij vier mbo-instellingen worden vergelijkbare focusgroepgesprekken gevoerd met vak- en taaldocenten in nog vijf andere opleidingen, en zal een breed opgezet interviewonderzoek worden gehouden onder een honderdtal werkgevers. De bij het project betrokken opleidingen gaan op basis van deze uitkomsten aan de slag met ontwerp van taalrijk en meer beroepsgericht onderwijs. De uitkomsten moeten leiden tot inzichten en conclusies die deze opleidingen verder helpen, maar die ze daarnaast ook overstijgen. Zo wordt kennis ontwikkeld over het belang van taalvaardigheidsontwikkeling in het mbo-onderwijs en hoe dat te doen. Ook worden daarmee mechanismen blootgelegd van de wijze waarop het onderwijs veranderende eisen van de beroepenvelden om kan zetten naar het onderwijs.

Bibliografie

Bawarshi, A.S., & Reiff, M.J. (2010). *Genre: An Introduction to History, Theory, Research, and Pedagogy*. West Lafayette, IN: Parlor Press/WAC Clearinghouse.

- Beaufort, A. (2009). Preparing Adolescents for the Literacy Demands of the 21st Century Workplace. In L. Christenbury, R. Bomer & P. Smagorinsky (Red.), *Handbook of Adolescent Literacy Research* (pp. 239-255). New York, NY: Guilford Press.
- Biber, D., & Conrad, S. (2009). *Register, Genre, and Style*. Cambridge: Cambridge University Press.
- Bol, T., & Rözer, J. (2017). Vakmanschap op de arbeidsmarkt. Een analyse van de arbeidsmarktuitkomsten van afgestudeerden uit het middelbaar beroepsonderwijs. *Mens en Maatschappij*, 92(3), 233-257.
- Bolle, T. (2009). *Drieslag Taal. Praktijkboek Taalbeleid Nederlands in het mbo*. Amsterdam: ITTA.
- Bolle, T. (2011). Ligt de drieslag taal' op koers? Ontwikkelingen van het Nederlands in het mbo. *Levende Talen Magazine*, 98(8), 16-20.
- Bolle, T. (2013). Op weg naar de examens Nederlands in het mbo. Taalcoaches werken aan taalbeleid. *Levende Talen Magazine*, 100(1), 16-19.
- Bolle, T., & Van Meelis, I. (2014). Wie taalbewust lesgeeft maakt onderwijs aantrekkelijker. Taalgerichte vakdidactiek in praktische leersituaties. *Van twaalf tot achttien*, 24(5), 42-45.
- Brandt, D. (2009). *Literacy and Learning: Reflections on Writing, Reading, and Society*. San Francisco, CA: Jossey Bass
- Britton, J.N. (1993). *Language and Learning: The Importance of Speech in Children's Development* (2e druk). Portsmouth, NH: Heinemann.
- Bruin, A. (2018). *Werk maken van taal in het mbo. Een onderzoek naar visies van mbo-docenten op de taaleisen die de dynamische arbeidsmarkt stelt aan mbo-studenten en hun opvattingen over de feitelijke en wenselijke aansluiting van het curriculum op die taaleisen*. (Ongepubliceerde masterscriptie). Rijksuniversiteit Groningen, Groningen.
- Buisman, M., & Van der Velden, R. (2017). *De toekomst van vakmanschap*. Amsterdam/Maastricht: Kohnstamm Instituut/Researchcentrum voor Onderwijs en Arbeidsmarkt.
- Christoffels, I., & Baay, P. (2016). *De toekomst begint vandaag: 21ste-eeuwse vaardigheden in het beroepsonderwijs*.s-Hertogenbosch: Expertisecentrum Beroepsonderwijs.
- Elbers, E. (2012). *Iedere les een taalles? Taalvaardigheid en vakonderwijs in het (v)mbo. De stand van zaken in theorie en onderzoek*. Utrecht/Den Haag: Universiteit Utrecht/PROO.
- Faber-De Lange, B., & Van der Pool, E. (2015). De communicatief competente professional in de ogen van managers. *Tijdschrift Voor Taalbeheersing*, 37(2), 217-257.
- Glopper, C.M. de (2016). *Projectaanvraag Taalcompetenties voor een dynamische arbeidsmarkt* (NRO- Aanvraagformulier Beroepsonderwijs 2017). Groningen: Rijksuniversiteit Groningen.
- Hajer, M. (2005). Taalgericht vakonderwijs: Tijd voor een nieuw vijfjarenplan. 3-11.
- Hajer, M., & Meestringa, T. (2004). *Handboek Taalgericht Vakonderwijs*. Bussum: Couthino.
- Hendrix, T., Hovens, T., & Kappers, A. (2012). Ten geleide. Taalonderwijs in het mbo – ontwikkelingen, kwesties, problemen en oplossingen. *Levende Talen Magazine*, 99(special), 4-5.
- Klein, P.D. (1999). Reopening Inquiry into Cognitive Processes in Writing-To-Learn. *Educational Psychology Review*, 11(3), 203-270.
- Lindlof, T.R., & Taylor, B.C. (2011). *Qualitative Communication Research Methods* (3^e druk). Thousand Oaks, CA: SAGE.
- Meng, C., Van der Meijden, A., Levels, M., Van Schooten, E., & Van der Velden, R. (2017). *Competency-oriented secondary VET-Effects on competencies, further education and labor market*. Amsterdam/Maastricht: Kohnstamm Instituut/Researchcentrum voor Onderwijs en Arbeidsmarkt.
- Mercer, N. (1995). *The Guided Construction of Knowledge: Talk Amongst Teachers and Learners*. Bristol: Multilingual Matters.
- Neuvel, J., Bersee, T., Den Exter, H., & Tijssen, M. (2004). *Nederlands in het middelbaar beroepsonderwijs. Een verkennend onderzoek naar het onderwijsaanbod Nederlands en de taalvaardigheid van de leerlingen*.s-Hertogenbosch: CINOP.

- Onderwijs in Cijfers (Ministerie van Onderwijs, Cultuur en Wetenschap, Dienst Uitvoering Onderwijs & Centraal Bureau voor de Statistiek)(z.d.). *Aansluiting mbo-arbeidsmarkt. Oordeel leerbedrijven over in dienst genomen mbo'ers*. Geraadpleegd op 8 mei 2018, <https://www.onderwijsincijfers.nl/kengetallen/mbo/aansluiting-mbo-arbeidsmarkt/oordeel-leerbedrijven-over-in-dienst-genomen-mbo>
- Onderwijsraad (2006). *Versteviging van kennis in het onderwijs*. Den Haag: Onderwijsraad.
- Raaphorst, E., & Steehouder, P. (2010). *Nederlandse taal in het mbo: beleid en praktijk*. 's-Hertogenbosch/Utrecht: Expertisecentrum Beroepsonderwijs.
- Raaphorst, E., & Steehouder, P. (2011). *Taal terug op het mbo. Beleid en praktijk van (nieuw) taalonderwijs in het middelbaar beroepsonderwijs*. 's-Hertogenbosch/Utrecht: Expertisecentrum Beroepsonderwijs.
- Turkenburg, M., & Vogels, R. (2017). *Beroep op het mbo*. Den Haag: Sociaal en Cultureel Planbureau.
- Van der Meulen, N. (2018). *Taalexamining van Nederlands op het mbo*. (Ongepubliceerde masterscriptie). Rijksuniversiteit Groningen, Groningen.
- Van Knippenberg, M.A.J. (2010). *Nederlands in het middelbaar beroepsonderwijs. Een casestudy in de opleiding Helpende Zorg*. Delft/Zutphen: Eburon.
- Van Knippenberg, M.A.J. (2012). Nederlands in het mbo. *Levende Talen Magazine*, 99, 6-11.

Is translanguaging een duurzame strategie voor het hoger onderwijs in Zuid-Afrika?

Adelia Carstens
Universiteit van Pretoria

TVT 41 (1): 41–54

DOI: 10.5117/TVT2019.1.003.CARS

Abstract

Is translanguaging a sustainable strategy in South African higher education?

Translanguaging could potentially harness content and language learning in a second or foreign language. However, there is a paucity of research on the role of language attitudes and language repertoires in determining translanguaging spaces in higher education. The research questions that guided this study were ‘What are learners’ perceptions about translanguaging in higher education?’, ‘What are the roles of learners’ and lecturers’ language repertoires in determining translanguaging spaces?’ and ‘How do these issues influence decisions on curriculum design?’ I first outline a framework that distinguishes between learner-directed and teacher-directed translanguaging. This is followed by an overview of two case studies on eliciting students’ appreciation of translanguaging. It is concluded that students are generally appreciative of translanguaging, although students who share their L1 with the lecturer are more positive than those who do not. Evidently, however, curricula that utilise translanguaging as a teaching and learning strategy should provide equal opportunities to speakers of all L1’s to succeed in their studies.

Keywords: higher education, multilingualism, translanguaging, translanguaging spaces, teacher-directed translanguaging, learner-directed translanguaging

Inleiding

In dit artikel staan twee casestudy's centraal die ik heb uitgevoerd aan een Zuid-Afrikaanse universiteit. In deze casestudy's werden de percepties van studenten onderzocht van *translanguaging* als een meertalige leerstrategie (zie hieronder). Op basis van de uitkomsten van deze casestudy's bespreek ik hoe het succes van *translanguaging* wordt beïnvloed door verschillen in de verbale repertoires van docenten en leerders. Aan het eind van dit artikel ga ik na hoe de verworven inzichten van belang kunnen zijn voor de mogelijke invoering van *translanguaging* in academische-geletterdheidscurricula van universiteiten in Zuid-Afrika met linguïstisch diverse studentenpopulaties maar Engels als de enige taal waarin onderwijs gegeven wordt.

Achtergrond

Wereldwijd is het Engels uitgegroeid tot de dominante taal van handel, wetenschap en technologie (Mazak, 2017, p. 6; Coleman, 2006, p. 2), met als gevolg een voorkeur voor Engels als dominante taal in het onderwijs (Van der Walt, 2013; Doiz, Lasagabaster, & Sierra, 2014; Mazak & Carroll, 2017). In ontwikkelende landen wordt het Engels bij voorkeur als de *lingua academia* gebruikt, omdat Engels wereldwijd academische en carrière-mobiliteit vergemakkelijkt. De dominantie van het Engels in het hoger onderwijs draagt echter ook bij aan de problemen waarvoor veel studenten zich gesteld zien (Kagwesage, 2013, p. 1; Dalvit & De Klerk, 2005; Weideman, 2006; Deyi, Simon, Ncobo, & Thole, 2007).

Een strategie die bevorderlijk is gebleken voor de toegankelijkheid van academisch onderwijs voor studenten voor wie het Engels een tweede of derde taal is, is hen toe te staan om in de communicatie met docenten en andere studenten gebruik te maken van hun moedertaal (Chan, 2015; Coyle, Hood, & Marsh, 2010; Darayi-Hansen, Barfod, & Schwarz, 2017; Lo & Lin, 2015; Mazak, Mendoza, & Mangonéz, 2017; Tian & Macraro, 2012). Deze strategie, die bekend is geworden als *translanguaging*, houdt in dat er ten minste twee talen tegelijk worden gebruikt om 'betekenis te geven, ervaringen te vormen en begrip en kennis te verwerven' (Baker, 2011, p. 288).¹ Zoals opgemerkt door Mazak (2017, p. 6) is de literatuur over *translanguaging* in het hoger onderwijs echter schaars.

Context

Het grootste deel van de 26 universiteiten in Zuid-Afrika heeft intussen gekozen voor Engels als de enige taal waarin onderwijs gegeven wordt (Madiba, 2018, p. 506). Veel moedertaalsprekers van inheemse Afrikatalen als isiZulu en Sepedi vinden onderwijs in het Engels lastig, onder meer omdat ze in het voortgezet onderwijs docenten hebben gehad die zelf niet bedreven waren in academisch Engels, en die het als gevolg van tekortkomingen in hun opleiding tijdens de Apartheid wellicht ook ontbrak aan conceptuele kennis op een hoog niveau (Nel & Müller, 2010). Bovendien worden op veel Engelstalige middelbare scholen die deze studenten hebben bezocht in de informele communicatie tussen leraren en studenten vaak ook andere talen dan Engels gebruikt, om zo de kloof te overbruggen tussen het Engels en de moedertaal van de studenten (Hurst, Madiba, & Moreira, 2017). Daarentegen hebben de meeste moedertaalsprekers van (het aan het Nederlands verwante) Afrikaans onderwijs gehad van docenten die ook Afrikaans als moedertaal hadden en daarnaast ook vloeiend Engels spraken, en die een inhoudelijk adequate opleiding genoten hadden. Hoewel Afrikaanssprekende studenten aldus over het algemeen beschikken over een solide inhoudelijke basis voor de verschillende academische disciplines, hebben ook sommigen van hen moeite met de wetenschappelijke registers van het Engels.

Helaas is de demografie van het onderwijzend personeel in Zuid-Afrikaanse universiteiten nog altijd geen afspiegeling van de demografie van de studentenpopulatie (Govender, Zondo, & Makgoba, 2013). Die situatie heeft een negatieve invloed op de flexibiliteit van *translanguaging* in academische geletterdheid en inhoudelijke vakken.

Op de Universiteit van Pretoria is het onderwijs in academische geletterdheid erop gericht studenten te helpen discipline-specifiek Engels te gaan beheersen. In colleges waarin academische geletterdheid centraal staat, wordt een variëteit aan pedagogische strategieën gebruikt, waaronder *translanguaging*. Daarnaast kunnen in dit onderwijs de (taal)voorkeuren van studenten worden achterhaald, kan worden geconstateerd wat in de praktijk het beste werkt, kan worden bijgedragen aan theorievorming op dit gebied, en kan een meer kritische houding worden bevorderd tegenover diverse pedagogische benaderingen (Hamman, 2018, p. 22).

Hieronder bespreek ik kort twee modellen die bij *translanguaging* worden gehanteerd.

Modellen van *translanguaging*

Lewis (2008) maakt onderscheid tussen docentgerichte *translanguaging* en student-gerichte *translanguaging*. Docentgerichte *translanguaging* is een geplande en gestructureerde activiteit door de docent (cf. Lewis, Jones, & Baker, 2012, p. 655), die functies omvat als herformulering van inhoud, instructie voor taalverwerving (om talige ontwikkeling te ondersteunen), instructie voor de verwerving van nieuwe kennis, en facilitering van het onderwijsleerproces. Docentgerichte *translanguaging* vereist dat de docent vertrouwd is met de moedertaal van de studenten (Fennema-Bloom, 2009/2010).

Studentgerichte *translanguaging* betreft het gebruik van *translanguaging* door meertalige studenten zonder veel ondersteuning van een facilitator (Lewis et al., 2012, p. 665). De eerste stap voor docenten is het creëren van een veilige ruimte voor studenten om met *translanguaging* te oefenen (Canagarajah, 2011, p. 415; Bhabha, 1994). De studenten kunnen dan passende activiteiten voor *translanguaging* kiezen, zoals het zoeken en gebruiken van informatie, het bespreken van inhoud (in hun moedertaal *L1*, en hun tweede taal *L2*), het maken van samenvattingen (in *L1* of *L2*) en het maken van een schrijfpodracht (meestal in *L2*). Studentgerichte *translanguaging* vereist geen twee- of meertaligheid van de docent.

Translanguaging moet niet worden verward met vertalen. Lewis et al. (2012, p. 659) delen vertaalactiviteiten als volgt in: vertaling van de leerstof door de docent voor de hele groep (bijvoorbeeld in de vorm van PowerPoint-dia's), vertaling van aspecten van de leerstof voor individuele studenten, en vertaling van vakgerelateerde terminologie.

Om mogelijkheden en problemen in het gebruik van *translanguaging* in het hoger onderwijs te illustreren bespreek ik hieronder twee verkennende casestudy's.

In casestudy 1, uitgevoerd in 2015, was de onderzoeksvraag hoe studenten met uiteenlopende moedertalen (Afrikaans, isiXhosa, isiZulu, Sepedi, Setswana, Xitsonga en Tshivenda) dachten over het nut van *translanguaging* voor (1) hun begrip van vakinhoudelijke concepten, (2) hun beheersing van de *L2* (hier: Engels) als academische taal, en (3) de ontwikkeling van de *L1* als academische taal. Er namen twee groepen deel: studenten die in hun opleiding (Bouwkunde) bij voorkeur onderwijs in het Engels volgden, en studenten die voor het onderwijs in diezelfde opleiding het Afrikaans prefereerden. Beide groepen volgden een college academische geletterdheid, dat in de eerste groep verzorgd werd door een van mijn collega's en

in de tweede groep door mijzelf. In beide groepen werd het college in het Engels gegeven, en omdat geen van beide docenten vertrouwd was met de Afrikatalen, werd in beide groepen alleen gebruik gemaakt van studentgerichte *translanguaging* (zie ook Carstens, 2016).

Casestudy 2, uitgevoerd in 2016, had tot doel meer inzicht te krijgen in de pedagogische waarde van studentgerichte en docentgerichte *translanguaging*, in dit geval voor (1) het begrip van vakinhoudelijke concepten, en (2) de beheersing van de L2 als academische taal. Omdat docentgerichte *translanguaging* vereist dat de docent ook de taal van de studenten beheerst, was bij deze casestudy maar één groep studenten betrokken: moedertaalsprekers van het Afrikaans (zie ook Carstens, 2017).

Casestudies

Casestudy 1

In 2015 onderzocht ik het gebruik van *translanguaging* als strategie om meertalige studenten die het vak constructie-economie volgden te ondersteunen.

Een van taken in het vak academische geletterdheid, is het begrijpend leren lezen van een academische tekst, in dit geval over afvalbeheer in de constructiefase van een project (McDonald & Smithers, 1998). *Translanguaging*, vertalen en ook visualisatie van informatie werden ingezet om studenten te helpen bij het gezamenlijk begrijpen van de concepten.

De studenten werden verdeeld in twee groepen op basis van de onderwijstaal waar ze voor kozen: Engels (n = 57) of Afrikaans (n = 55). De Afrikaans-groep was homogeen in termen van de L1: alle studenten waren L1-sprekers van het Afrikaans. Deze studenten werd verzocht de groep in subgroepen van ongeveer tien studenten te verdelen. De groep die Engels had gekozen, werd gevraagd om subgroepen te vormen op basis van hun L1. Hier werden subgroepen gevormd van negen isiXhosa-, zeven isiZulu-, tien Sepedi-, zeven Setswana-, vier Xitsonga-, vier Tshivenda- en zestien Engelssprekenden. In beide groepen werd aan elke subgroep gevraagd om open plekken in te vullen op een kaart waarop de belangrijkste concepten en sub-concepten op het gebied van afvalbeheer al van Engelstalige labels waren voorzien. Gevraagd werd (1) bij de ingevulde termen uit een willekeurig gerangschikte lijst de Engelstalige definities te zoeken, (2) de termen uit het Engels naar de eigen L1 te vertalen, en (3) de definities van die termen naar hun eigen L1 te vertalen. De diverse taalspecifiek samengestelde

subgroepen werden gestimuleerd om bij het bespreken en uitvoeren van de opdracht zowel Engels als hun eigen L1 te gebruiken. Moedertaalsprekers van het Engels hoefden niet mee te doen aan de activiteiten (2) en (3). Dat bracht het totaal aantal studenten in deze groep dat daar wel aan deelnam op 41.

Nadat de studenten in beide groepen deze taak hadden voltooid, beantwoordden ze een aantal vragen in een semigestructureerde vragenlijst. De vragen waren gericht op hun percepties van *translanguaging* als ondersteuning van betekenisconstructie, van L2-ontwikkeling, en van terminologiecreatie in de L1. De antwoorden werden gecodeerd in drie cycli, met gebruikmaking van het kwalitatieve inhoudsanalyseprogramma Atlas.ti. In Tabel 1 hieronder worden voor beide groepen samen de toegewezen codes en de thema's waarin de antwoorden van studenten werden geclusterd gepresenteerd, evenals de saillante motivaties die ze daarbij gaven.

Zesentachtig procent van de sprekers van Afrikatalen en 91% van de sprekers van het Afrikaans liet weten dat *translanguaging* hen had geholpen om de concepten beter te begrijpen. Hetzelfde percentage van sprekers van Afrikaans (91%) meende dat de ontwikkeling van hun zwakkere taal (Engels) ook was bevorderd door het gebruik van *translanguaging*, terwijl slechts 51% van de sprekers van Afrikatalen dat als een voordeel van *translanguaging* had gezien. Drieënzestig procent van de sprekers van Afrikatalen was positief over de rol van *translanguaging* bij het creëren van nieuwe termen (L1-ontwikkeling).

Alles bijeen vonden de studenten *translanguaging* een nuttige strategie. Wel stonden de sprekers van het Afrikaans positiever tegenover *translanguaging* dan de sprekers van de Afrikatalen. Voor beide groepen golden de cognitieve effecten als het meest prominente voordeel. Bijna alle sprekers van het Afrikaans en ongeveer een derde van de sprekers van Afrikatalen vonden dat *translanguaging* hen had geholpen het concept van afvalbeheer beter te begrijpen. Alleen de isiXhosa L1-groep liet weten dat *translanguaging* bij hen het begrip niet had ondersteund, omdat hun moedertaal daarvoor 'te complex' was. Deze uitkomst sluit aan bij observaties van andere onderzoekers (Dyers & Davids, 2015, p. 21), die vonden dat sprekers van isiXhosa liever Engels gebruiken aangezien isiXhosa wordt beschouwd als 'moeilijk', 'ingewikkeld' en 'diep', mogelijk omdat ze isiXhosa nooit als een academische taal hebben gebruikt, en ook omdat standaard isiXhosa wordt gezien als literair en poëtisch.

In beide groepen werden daarnaast voordelen van affectieve aard vermeld. Moedertaalsprekers van Afrikatalen benadrukten vooral dat met *translanguaging* een veilige omgeving wordt gecreëerd om de mogelijkheden

Tabel 1 Overzicht van uitkomsten van casestudy 1 (Carstens, 2016)

Thema 1:	Thema 2:	Thema 3:	
Betekenisconstructie (Vragen 1 and 2)	L2-ontwikkeling (Vraag 3)	L1-ontwikkeling (terminologiecreatie) (Vragen 4 and 5)	
Code 1 <i>Translanguaging</i> hielp studenten het concept afvalbeheer te begrijpen, en aldus:	Code 3 <i>Translanguaging</i> hielp studenten bij de ontwikkeling van hun competentie en zelfvertrouwen in de L2, vanwege:	Code 5 <i>Translanguaging</i> kan dienen als platform om termen te scheppen in Afrikatalen, vanwege de voordelen in de:	Code 7 Intentie om in de toekomst L1-terminen te gebruiken, om:
* <i>de grote lijn te doorzien</i>	* <i>het uitbreiden van de woordenschat in het Engels</i>	* <i>communicatie (ervoor zorgen dat je begrepen wordt)</i>	* <i>de sociale cohesie te bevorderen onder sprekers van verschillende Afrikatalen</i>
* <i>onderscheid te maken tussen verschillende concepten</i>	* <i>het vergroten van hun zelfvertrouwen bij het gebruik van het Engels</i>	* <i>praktijk (in concrete situaties met de termen handelen)</i>	* <i>de sociale cohesie te bevorderen onder leden van dezelfde taalgemeenschap</i>
* <i>complexe concepten te vereenvoudigen</i>			
* <i>eigen conceptuele inzichten tot uitdrukking te brengen</i>			
Code 2 <i>Translanguaging</i> hielp studenten NIET bij het begrijpen van het concept afvalbeheer, vanwege:	Code 4 <i>Translanguaging</i> hielp studenten NIET bij de ontwikkeling van hun competentie en zelfvertrouwen in de L2, vanwege:	Code 6 <i>Translanguaging</i> zal hoogstwaarschijnlijk NIET gaan dienen als platform voor het scheppen van termen in Afrikatalen, vanwege:	Code 8 GEEN Intentie om in de toekomst L1-terminen te gebruiken, vanwege:
* <i>dialectische variatie in de L1</i>	* <i>'Engels is mijn sterkste taal'</i>	* <i>complexiteit van de L1</i>	* <i>complexiteit van de L1</i>
* <i>complexiteit van de L1</i>		* <i>'Engels is de wereldtaal'</i>	* <i>dialectische variatie in de L1</i>
* <i>ontbreken van een gespecialiseerde woordenschat in L1</i>			
* <i>'Engels is mijn sterkste taal'</i>			
* <i>'Engels is de taal van wetenschappelijke communicatie'</i>			

te verkennen om in de L1 wetenschappelijke concepten te verwoorden. Moedertaalprekers van het Afrikaans wezen er vooral op dat ze zich konden uiten in de taal waarin ze het meest bedreven waren (het Afrikaans), en ze vonden het positief dat ze daarnaast een voor hen bekende taal (het Engels) konden gebruiken om met anderstaligen samen te werken.

Twee derde van de sprekers van Afrikatalen waardeerde de gelegenheid om in de moedertaal met terminologiecreatie te experimenteren, vooral vanwege de mogelijkheid die nieuwe termen bieden om inclusiviteit te bevorderen. In het oog springende redenen om maar gematigd positief te zijn over terminologiecreatie betroffen het veronderstelde tekort aan bestaande termen in de L1 en ook de complexiteit van de L1, die het nodig zou maken om ingewikkelde syntactische constructies te gebruiken.

Hoewel casestudy 1 laat zien dat *translanguaging* een nuttig instrument kan zijn voor het realiseren van diverse pedagogische functies in een meertalige context, besef ik dat een beperking in deze casestudy was dat ik als docent niet vertrouwd was met de moedertalen van de studenten in een van de twee groepen. Daardoor was ik niet goed in staat om de studenten in deze groep op hun eigen terrein in hun L1 tegemoet te komen en zo mijn cognitieve doelen (begrip vergroten) en mijn affectieve doelen (solidariteit uiten) te bereiken. Ik was alleen in staat om voor de studenten ruimte te creëren om onderling *translanguaging* te gebruiken.

Casestudy 2

Met casestudy 2 wilde ik – als spreker van zowel Afrikaans als Engels – meer inzicht krijgen in studentgerichte en docentgerichte vormen van *translanguaging*, in dit geval bij studenten die Afrikaans als moedertaal hadden.

Bij 31 moedertaalsprekers van het Afrikaans werden pre- en post-interventievragenlijsten afgenomen. In de pre-interventievragenlijst (pre-V) werd gevraagd naar de voorkeuren van de studenten voordat ze aan het onderwijs met *translanguaging* (vanaf hier: de interventie) deel gingen nemen, terwijl in de post-interventie vragenlijst (post-V) werd gevraagd naar de gang van zaken zoals ze die in de praktijk hadden ervaren.

De uitkomsten zijn samengevat in Tabel 2 en in Tabel 3 (alle genoemde getallen staan voor percentages).

Voorkeuren bij docentgerichte *translanguaging*. Voorafgaand aan de interventie gaf bijna de helft van de studenten er de voorkeur aan dat concepten werden uitgelegd in het Afrikaans en bijna de andere helft prefereerde een uitleg in beide talen. Een kleine minderheid koos voor Engels. De eerste groep noemde vertrouwdheid met het Afrikaans als academische taal als belangrijkste argument; in de tweede en de derde groep meenden veel

Tabel 2 Docentgerichte *translanguaging*

	Uitleg van concepten				Vergroten van woordenschat en grammaticale kennis				Organisatie van het college				Communiceren met individuele studenten			
	Afr	Eng	Beide		Afr	Eng	Beide		Afr	Eng	Beide		Afr	Eng	Beide	
Pre-V	48	4	48		29	26	45		77	0	23		90	0	10	
Post-V	100	0	0		100	0	0		100	0	0		100	0	0	

Tabel 3 Studentgerichte *translanguaging*

	Bespreken van concepten		Plannen van opdrachten		Maken van aantekeningen		Maken van literatuur-samenvattingen		Schrijven van eerste versie van een verslag		Schrijven van definitieve versie van een verslag	
	Afr	Eng	Afr	Eng	Afr	Eng	Afr	Eng	Afr	Eng	Afr	Eng
Pre-V	94	6	94	6	81	19	39	22	50	50	29	71
Post-V	100	0	100	0	100	0	26	71	0	100	0	100

studenten dat onderwijs dat ze (ook) in het Engels zouden volgen, de beste manier is om in die taal academisch geletterd te worden.

Tijdens de interventie bleek dat de behandelde bouwkundige concepten zo goed werden toegelicht in het bouwkundecollege dat ze niet opnieuw hoefden te worden besproken in het academischegeletterdheidscollege. De post-V-vragen over dit onderwerp waren achteraf gezien dan ook niet relevant. Ook lieten alle studenten weten dat het vergroten van de woordenschat en de grammaticale kennis in de praktijk overwegend in het Afrikaans plaats had gevonden.

Voor het vergroten van de woordenschat en de grammaticale kennis gaf bijna de helft van de studenten de voorkeur aan een combinatie van beide talen; de voorkeuren van de andere studenten waren hier gelijkelijk verdeeld over Afrikaans en Engels. Tijdens de interventie werd er in de praktijk zo veel in het Afrikaans gecommuniceerd dat de studenten in de post-V al lemaal de optie Afrikaans aankruisten.

Voorafgaand aan de interventie lieten de meeste studenten weten dat ze er de voorkeur aan gaven als de docent in het Afrikaans met hen communiceerde over de organisatie van het onderwijs en als de docent individuele studenten ook in het Afrikaans aansprak. Afgesproken werd dat tijdens de interventie voor de organisatie van het college en in de communicatie met individuele studenten Afrikaans gebruikt zou worden. Uit de antwoorden in de post-V blijkt dat dat inderdaad gebeurde.

Wat de docentgerichte *translanguaging* betreft lijkt het er dus op dat, hoewel de Afrikaanssprekende studenten goed tweetalig zijn wanneer zij op de universiteit arriveren, ze de voorkeur geven aan instructie in de academische code waarmee ze het meest vertrouwd zijn.

Voorkeuren bij studentgerichte translanguaging. Voorafgaand aan de interventie bleek een sterke voorkeur van de studenten voor het gebruik van de moedertaal in de discussies in kleine groepen. Na de interventie antwoordden de studenten dat ze in de praktijk inderdaad in de L1 met hun medestudenten hadden gecommuniceerd.

Voorafgaand aan de interventie bleek ongeveer 80% van de studenten er de voorkeur aan te geven tijdens het college aantekeningen in het Afrikaans te maken, terwijl slechts ongeveer 20% de voorkeur gaf aan het gebruik van het Engels voor dit doel. Na de interventie bleek dat alle studenten hun aantekeningen inderdaad in het Afrikaans hadden gemaakt. Dat lag ook voor de hand, aangezien het grootste deel van de leerstof door de docent in het Afrikaans werd overgedragen.

Voorafgaand aan de interventie waren de antwoorden over de literatuursamenvattingen gemengd: 39% had een voorkeur voor Afrikaans, 22% had een

voorkeur voor Engels, en 39% had geen voorkeur. Na de interventie bleek dat de meeste studenten de samenvattingstaken in het Engels hadden uitgevoerd.

Voorafgaand aan de interventie gaf de helft van de studenten te kennen dat ze de eerste versie van een groepsverslag het liefst in het Afrikaans schreven, terwijl minder dan een derde van de studenten de definitieve versie van het verslag bij voorkeur in het Afrikaans schreef. Uiteindelijk schreven alle groepen op één na de eerste versie in het Engels, en schreven alle groepen de definitieve versie in het Engels.

Deze bevindingen duiden erop dat, hoewel de Afrikaanssprekende studenten aanvankelijk terughoudend zijn om in de klas academisch Engels te spreken, ze al snel ontdekken dat het voordelig is om van meet af aan in het Engels te gaan schrijven. Gelet op de sterk toenemende tendens in het Zuid-Afrikaanse academisch onderwijs om de studenten (dus ook de Afrikaanstalige) ongeacht hun moedertaal alle schrijftaken in het Engels uit te laten voeren, is deze uitkomst niet verwonderlijk.

Conclusies en reflectie

In casestudy 1 bleken alle L2-sprekers van het Engels *translanguaging* positief te beoordelen. Moedertaalsprekers van Afrikatalen waren echter minder positief dan moedertaalsprekers van het Afrikaans. Hoogstwaarschijnlijk komt dat doordat veel sprekers van Afrikatalen hun moedertaal nooit als een academische taal hebben gebruikt en daarom Engels als hun sterkste taal beschouwen, vooral in academische communicatie. Bovendien was ik als docent met geen van de Afrikatalen vertrouwd. Hoewel pleitbezorgers van *translanguaging* de flexibiliteit van deze strategie benadrukken (zie de metafoor van een 'all-terrain vehicle' in Garcia, 2011), hebben mijn ervaringen in deze twee casestudy's me geleerd dat verschillen in het taalrepertoire van docenten en studenten de mogelijkheden om *translanguaging* in te zetten wel degelijk negatief beïnvloeden.

In beide casestudy's bleken moedertaalsprekers van het Afrikaans *translanguaging* te ervaren als een strategie die helpt de kloof tussen middelbare school en universiteit te overbruggen. Hun sterke identificatie met het Afrikaans als de taal waarin ze 'geboren' zijn, gecombineerd met de mogelijkheden voor zelfexpressie in een vertrouwde taal droegen bij aan hun positieve beoordeling van de *translanguaging*-strategie.

Duidelijk werd dat de voordelen van *translanguaging* groter zijn als behalve studentgerichte *translanguaging* ook docentgerichte *translanguaging* mogelijk is; dat is alleen het geval als de docent ook vertrouwd is met de moedertaal van de studenten.

Duidelijk is echter ook dat de doelstelling van onderling goed op elkaar afgestemde talige repertoires van docenten en studenten in veeltalige universiteiten op de korte of middellange termijn niet gemakkelijk te bereiken is. Dat is des te relevanter nu duidelijk is dat de linguïstische diversiteit in studentpopulaties alleen maar zal toenemen en de ‘meertalige universiteit’ mondiaal gezien de norm wordt. Gegeven dat het onderwijsbeleid van Zuid-Afrikaanse universiteiten meer en meer eentalig wordt (alleen nog Engels namelijk), is het des te belangrijker dat voor studenten met alle mogelijke L1's gelijke kansen worden gecreëerd. In deze context is het de taak van academische-geletterheidsdocenten om alle studenten te helpen hun studie succesvol af te ronden, terwijl ze gebruik maken van “the languages they know and the languages they are getting to know” (Van der Walt, 2013, p. 12). Voor dit doel kan het soort opdrachten geschikt zijn die in casestudy 1 gebruikt werden. Aangepast studiemateriaal, zoals meertalige woordenlijsten met definities in alle officiële landstalen, kan bijdragen aan conceptgeletterdheid en kan tegelijk de standaardisering versnellen van terminologie in talen die tot nu toe als niet meer dan informele omgangstalen beschouwd worden. Met een dergelijke pedagogische aanpak kan *translanguaging*, om het in de woorden van Garcia en Li (2014) te zeggen, opbloeien “in the cracks and crevices of communication with others who language differently, gradually becoming in and of itself a way of languaging through complex communicative interactions”.

Noot

Translanguaging moet niet worden verward met ‘receptieve meertaligheid’ (ook wel: communiceren in een ‘lingua receptiva’ of ‘luistertaal’). Receptieve meertaligheid kan worden ingezet in contexten waarin de ene deelnemer (of groep van deelnemers) aan de communicatie genoeg begrijpt van wat de andere deelnemer (of groep van deelnemers) in diens eigen taal naar voren brengt. Zo kunnen een moedertaalspreker van het Nederlands en een moedertaalspreker van het Afrikaans receptieve meertaligheid inzetten om een zinvol gesprek te voeren zonder dat ze zich in elkaars taal hoeven te kunnen uiten. Anders dan bij (studentgerichte) *translanguaging* is het voor receptieve meertaligheid noodzakelijk dat de communicatiedeelnemers elkaars taal tenminste enigermate beheersen (zie verder Blees & Ten Thije, 2017, p. 335).

Bibliografie

Baker, C. (Ed.) (2011). *Foundations of Bilingual Education and Bilingualism: A Social Approach*. (5th ed.). Bristol: Multilingual Matters.

- Bhabha, H. K. (1994). *The Location of Culture*. London: Routledge.
- Blees, G. J., & Ten Thije, J. D. (2017). Receptive multilingualism and awareness. In J. Cenoz & D. Gorter (Eds.), *Language Awareness and Multilingualism: Encyclopedia of Language and Education* (3rd ed.) (pp. 333-345). Cham: Springer International Publishing.
- Canagarajah, S. (2011). Codemeshing in academic writing: Identifying teachable strategies in translanguaging. *The Modern Language Journal*, 95(3), 401-417.
- Carstens, A. (2016). Translanguaging as a vehicle for L2 acquisition and L1 development: students' perceptions. *Language Matters*, 47(2), 203-222.
- Carstens, A. (2017). Die wisselwerking tussen teorie en praktyk in die ontwerp van 'n kurrikulum met transtaling as kernkomponent. *Stellenbosch Papers in Linguistics Plus*, 53, 25-39.
- Chan, S. (2015). Linguistic challenges in the mathematical register for EFL learners: Linguistic and multimodal strategies to help learners tackle mathematical word problems. *International Journal of Bilingual Education and Bilingualism*, 18(3), 306-318.
- Coleman, J. A. (2006). English-medium teaching in European higher education. *Language Teaching*, 39, 1-14.
- Coyle, D., Hood, P., & Marsh, D. (2010). *CLIL: Content and Language Integrated Learning*. Cambridge: Cambridge University Press.
- Cummins, J. (1979). Cognitive/academic language proficiency, linguistic interdependence, the optimum age question, and some other matters. *Working Papers on Bilingualism*, 19, 121-129.
- Dalvit, L., & De Klerk, V. (2005). Attitudes of Xhosa-speaking students at the University of Fort Hare towards the use of Xhosa as a language of learning and teaching (LOLT). *Southern African Linguistics and Applied Language Studies*, 23(1), 1-18.
- Darayi-Hansen, P., Barfod, S., & Schwarz, L. (2017). A call for (trans)languaging: The language profiles at Roskilde University. In C.M. Mazak & K.S. Carroll (Eds.), *Beyond Monolingual Ideologies. Translanguaging in Higher Education* (pp. 29-49). Bristol: Multilingual Matters.
- Deyi, S., Simon, E., Ncobo S., & Thole, A. (2007). Promoting the multilingual classroom: Why the significance of multilingualism in higher education? *Paradigms*, 14, 10-20.
- Doiz, A., Lasagabaster, D., & Sierra, J.M. (Eds.). (2014). *English medium instruction at universities: Global challenges*. Bristol: Multilingual Matters.
- Dyers, C., & Davids, G. (2015). Post-modern 'languagers': The effects of texting by university students on three South African languages. *Southern African Linguistics and Applied Language Studies*, 33(1), 21-30.
- Fennema-Bloom, J. R. (2009/2010). Code-scaffolding: A pedagogic code-switching technique for bilingual content instruction. *Journal of Education*, 190(3), 27-34.
- Garcia, O. (2011). Educating New York's bilingual children: Constructing a future from the past. *International Journal of Bilingual Education and Bilingualism*, 14, 133-153.
- Garcia, O., & Li, W. (2014). *Translanguaging, Language, Bilingualism and Education*. Basingstoke: Palgrave Macmillan.
- Govinder K. S., Zondo, N. P., & Makgoba, M. W. (2013). A new look at demographic transformation for universities in South Africa. *South African Journal of Science*, 109(11/12), 1-11.
- Hamman, L. (2018). Translanguaging and positioning in two-way dual language classrooms: A case study for criticality. *Language and Education*, 32(1), 21-42.
- Hurst, E., Madiba, M., & Morreira, S. (2017). Surfacing and valuing students' linguistic resources in an English-dominant university. In D.M. Palfreyman & C. Van der Walt (Eds.), *Academic Bilingualities: Multilingual Repertoires in Higher Education* (pp. 76-95). Bristol: Multilingual Matters.
- Kagwesage, A. M. (2013). Coping with English as language of instruction in higher education in Rwanda. *International Journal of Higher Education*, 2(2), 1-12.

- Lewis, G., Jones, B., & Baker, C. (2012). Translanguaging: Developing its conceptualisation and contextualisation. *Educational Research and Evaluation*, 18(7), 655-670.
- Lewis, W. G. (2008). Current challenges in bilingual education in Wales. Multilingualism and Minority Languages: Achievements and Challenges in Education. *AILA Review*, 21, 69-86.
- Madiba, M. 2018. The multilingual university. In A. Creese & A. Blackledge (Eds.), *Routledge Handbook of Language and Superdiversity* (pp. 504-517). London: Routledge.
- Mazak, C. M. (2017). Introduction: Theorizing translanguaging practices in higher education. In C. M. Mazak & K. S. Carroll (Eds.), *Beyond Monolingual Ideologies. Translanguaging in Higher Education* (pp. 1-10). Bristol: Multilingual Matters.
- Mazak, C. M., & Carroll, K. S. (Eds.). (2017). *Beyond Monolingual Ideologies. Translanguaging in Higher Education*. Bristol: Multilingual Matters.
- Mazak, C. M., Mendoza, F., & Mangonéz, L. P. (2017). Professors translanguaging in practice: Three cases from a bilingual university. In C. M. Mazak & K. S. Carroll (Eds.), *Beyond Monolingual Ideologies: Translanguaging in Higher Education* (pp. 70-90). Bristol: Multilingual Matters.
- McDonald, B. & Smithers, M. (1998). Implementing a waste management plan during the construction phase of a project: A case study. *Construction Management and Economics* 16(1), 71-78.
- Nel, N., & Müller, H. (2010). The impact of teachers' limited English proficiency on English second language learners in South African schools. *South African Journal of Education*, 30(4), 635-650.
- Tian, L., & Macaro, E. (2012). Comparing the effect of teacher codeswitching with English-only explanations on the vocabulary acquisition of Chinese university students: A lexical focus-on-form study. *Language Teaching Research*, 16(3), 367-391.
- Van der Walt, C. (2013). *Multilingual Higher Education: Beyond English Medium Orientations*. Bristol: Multilingual Matters.
- Weideman, A. (2006). Transparency and accountability in applied linguistics. *Southern African Linguistics and Applied Language Studies*, 24(1), 71-86.

En dan zit je met de gebakken peren!

Strategisch manoeuvreren met het argumentum ad consequentiam

Frans van Eemeren en Bart Garssen
Universiteit van Amsterdam en Ilias

TVT 41 (1): 55–68

DOI: 10.5117/TVT2019.1.004.EEME

Abstract

You have to face to consequences! Strategic maneuvering with the argumentum ad consequentiam

In argumentative discourse fallacies occur regularly, but often they seem not to be noticed by the participants. This also goes for the fallacy known as the *argumentum ad consequentiam*, in spite of the fact that people generally denounce this fallacy when confronted with clear cases. In this paper it is argued that the explanation of this paradox is that the different types of *argumentum ad consequentiam* are modes of strategic maneuvering which take on a reasonable appearance when they mimic legitimate pragmatic argumentation or *reductio ad absurdum* argumentation. The following hypothesis was tested in an experiment: instances of the *argumentum ad consequentiam* are regarded less unreasonable than clear cases if they are presented as reasonable pragmatic argumentation.

Keywords: *argumentum ad consequentiam*, fallacies, hidden fallaciousness, pragma-dialectics, pragmatic argumentation, strategic maneuvering

Inleiding

In een discussie met evolutionair bioloog en fervent atheïst Richard Dawkins verdedigt Wendy Wright, voorzitter van de conservatief-christelijke Concerned Women of America, het creationisme. Ze ondersteunt haar standpunt dat er een liefdevolle “schepper” bestaat als volgt: “What

a person believes about how human beings are created, shapes what they believe about human beings. And if we believe that human beings were created out of love, that is by a loving creator, and [that he] has given each one of us not only a material body but a spirit and a soul we are more likely to treat others with respect and dignity”.¹ Kort gezegd: er bestaat een liefdevolle schepper want als we daarin geloven zullen we anderen met meer respect en waardigheid behandelen. Zo tracht Wright tracht te bewijzen dat iets bestaat door te wijzen op de gunstige gevolgen van die existentie.

Iets dergelijks gebeurde ook nadat de Amerikaanse rapper Kanye West op 2 mei 2018 in een interview met de Amerikaanse entertainmentzender TMZ gesteld had dat slavernij wat hem betreft een keuze is: “als je hoort dat slavernij 400 jaar duurde. 400 jaar? Dat klinkt als een keuze”. Op deze woorden werd hij door TMZ-medewerker Van Lathan als volgt aangesproken: “Je mag denken wat je wil, maar jouw uitspraken hebben wereldwijde gevolgen. Terwijl jij muziek maakt en het luxeleven van een artiest leidt, worden wij – de rest van de samenleving – nog steeds gemarginaliseerd door de gevolgen van 400 jaar slavernij”.² Vervolgens herhaalde West dat slavernij een keuze is.

Zowel Wendy Wright als Van Lathan wijzen op de gevolgen of consequenties van een bewering over een bepaalde stand van zaken (‘er bestaat een liefhebbende “schepper”’ respectievelijk ‘slavernij is geen keuze’). Ze baseren hun argumentatie op de positieve of negatieve gevolgen van wat er in het standpunt wordt beweerd. Dit maakt dat er in beide gevallen sprake is van de drogreden die bekend staat als het *argumentum ad consequentiam*. In een *argumentum ad consequentiam* wordt een niet-gelegitimeerde stap gezet van een normatieve premisse naar een descriptief standpunt. Een bewering over het al dan niet bestaan van een bepaalde stand van zaken wordt verdedigd door te wijzen op de positieve respectievelijk negatieve gevolgen die deze stand van zaken zou hebben.

Het *argumentum ad consequentiam* is niet een drogreden die we dagelijks overal tegenkomen. Deze onredelijke discussie-zet lijkt wel geregeld voor te komen in discussies over religie en discriminatie. Met name in situaties waarin sprekers hun toevlucht zoeken tot *wishful thinking* valt de *ad consequentiam*-drogreden vaak te beluisteren. De Britse WOII-historicus Anthony Beevor beschrijft in deze zin de wanhopige uitspraken van Duitse soldaten na hun nederlaag bij Stalingrad:

Soviet intelligence officers, however, found their German prisoners still in a state of denial and confused logic at the possibility of defeat. ‘We have got to believe that Germany will win the war,’ said a Luftwaffe navigator from a JU 52 shot down on the Stalingrad run, ‘or what is the use of going on with it?’

A soldier reflecting the same obstinacy: ‘if we lose the war we have nothing to hope for.’ (Beevor, 2012, p. 257)

In dit soort eenvoudige gevallen is het onredelijke karakter van deze argumentatieve zet meteen duidelijk, ook voor leken die nog nooit van deze drogredes hebben gehoord. Dit blijkt uit ons empirisch onderzoek naar redelijkheidsoordelen van gewone taalgebruikers (Van Eemeren, Garssen, & Meuffels, 2009). De resultaten van dit experimentele onderzoek wijzen uit dat taalgebruikers het *argumentum ad consequentiam* als een zeer onredelijke drogreden beoordelen (pp. 176-179). In ons onderzoek gingen we uit van “clear cases”. In de dagelijkse praktijk van het argumenteren zullen de drogredes zich meestal niet zo herkenbaar manifesteren en is het onredelijke karakter ervan ook veel lastiger bloot te leggen. Daarom rijst de vraag hoe de argumentatie waarin het *argumentum ad consequentiam* voorkomt in de praktijk zo kan worden gepresenteerd dat deze drogreden een redelijker voorkomen krijgt. We zullen hiervoor een verklaring geven en die vervolgens toetsen aan de hand van een experiment.

Het onderzoek waarover we hier rapporteren maakt deel uit van het onderzoeksproject met de naam “hidden fallaciousness” (verborgen drogredelijkheid), waarin we nagaan we hoe taalgebruikers strategisch manoeuvreren om de drogredes die ze begaan een minder onredelijk voorkomen te geven. In eerder onderzoek in dit project zijn we nagegaan hoe het *argumentum ad hominem* (Van Eemeren, Garssen, & Meuffels, 2012) en het *argumentum ad baculum* (Van Eemeren, Garssen, & Meuffels, 2015) in de praktijk kunnen worden gecamoufleerd. In deze bijdrage zullen we eerst de pragma-dialectische opvatting het *argumentum ad consequentiam* uiteenzetten. Vervolgens schenken we aandacht aan het strategisch manoeuvreren met het *argumentum ad consequentiam*. Daarna bespreken we een experiment dat betrekking heeft op een van de verschijningsvormen van de *ad consequentiam*-drogreden.

Twee vormen van het *argumentum ad consequentiam*

In de pragma-dialectische argumentatietheorie worden drogredes opgevat als overtredingen van regels voor een kritische discussie. Het *argumentum ad consequentiam* vormt een overtreding van de argumentatieschema-regel (Regel 8): Een standpunt mag niet als afdoende verdedigd worden beschouwd als de verdediging niet plaatsvindt door middel van

een geschikt argumentatieschema dat correct is toegepast (Van Eemeren & Grootendorst, 2004, p. 194).

Het argumentatieschema, de centrale notie van deze regel, karakteriseert de manier waarop de aanvaardbaarheid van het argument wordt overgebracht op het standpunt. De argumentatieschemaregel bevat feitelijk twee onderdelen: 1) het argumentatieschema moet geschikt zijn en 2) het argumentatieschema moet correct zijn toegepast.

De geschiktheid van het argumentatieschema is in die zin een intersubjectieve aangelegenheid dat de protagonist en de antagonist in principe in de openingsfase van een kritische discussie overeenkomen (of geacht worden te zijn overeengekomen) welke argumentatieschema's in de lopende discussie mogen worden toegepast. Sommige argumentatieschema's zijn niet geschikt voor gebruik in bepaalde contexten of zijn zelfs per definitie niet geschikt. Dat laatste kan bijvoorbeeld te wijten zijn aan een inadequate combinatie van het type standpunt dat verdedigd wordt en het argument dat ter verdediging wordt aangevoerd. Dit is wat bijvoorbeeld het geval is bij het *argumentum ad consequentiam*, waarin een normatief argument een descriptief standpunt moet rechtvaardigen (Van Eemeren & Grootendorst, 1992, p. 162).

Of een geschikt bevonden argumentatieschema ook correct is toegepast, wordt in de toetsingsprocedure vastgesteld aan de hand van bij het argumentatieschema behorende kritische vragen van de antagonist en de reacties hierop van de protagonist. Uit het feit dat er bij het *argumentum ad consequentiam* geen bijbehorende kritische vragen te bedenken zijn, blijkt eigenlijk al dat we hier met een argumentatieschema te maken hebben dat inherent ongeschikt is als redelijk verdedigingsmiddel.

Het is de vraag hoe een vorm van argumentatie die er op het eerste gezicht zo onredelijk uitziet toch met enig succes kan worden toegepast in discussies. Waarom zouden discussianten zichzelf openlijk als onredelijk opstellen? Het antwoord op deze vraag kan worden gevonden in de met het strategisch manoeuvreren uitgebreide pragma-dialectiek (van Eemeren, 2010).

In hun streven naar effectiviteit zullen discussianten zodanig strategisch proberen te manoeuvreren dat ze door middel van hun discussiebijdragen – zich houdend aan de regels voor kritische discussie – hun dialectisch doel kunnen realiseren terwijl ze *tegelijkertijd* trachten hun retorisch doel te bereiken dat ze de discussie winnen door hun standpunt aanvaard te krijgen door de tegenpartij. Bij het balanceren tussen enerzijds dialectische oplossingsgerichtheid en anderzijds retorische effectiviteit om deze twee (soms haaks op elkaar staande) doelen tegelijkertijd te realiseren, maken discussianten gebruik van strategisch manoeuvreren (Van Eemeren, 2010, p. 40).

Op zichzelf is er niets mis met de behoefte van discussianten om een discussie te winnen, maar als deze behoefte de overhand krijgt dan kan dat tot ontsporingen leiden waarin het ongebreideld nastreven van effectiviteit resulteert in een ontsporing van het strategisch manoeuvreren en discussianten hun gebondenheden aan een redelijke gedachtewisseling opgeven. In deze optiek zijn drogredenen ontsporingen van strategisch manoeuvreren die een overtreding inhouden van de regels voor het voeren van een kritische discussie. Door de regels voor kritische discussie te schenden, blokkeren of frustreren zulke argumentatieve zetten het proces van het oplossen van een meningsverschil op redelijke gronden, zodat deze vormen van strategisch manoeuvreren als drogredelijk moeten worden gekarakteriseerd.

Ontsporingen van strategisch manoeuvreren kunnen gemakkelijk aan de aandacht van de discussianten ontsnappen omdat afwijkingen van de regels voor kritische discussie in de argumentatieve praktijk van alledag vaak moeilijk te ontdekken zijn: geen van de partijen in een discussie zal geneigd zal om zichzelf openlijk (“on record”) als een onredelijk iemand te afficheren – al was het maar om te voorkomen dat de betreffende discussiebijdrage daardoor totaal ineffectief zou worden:

so arguers will most likely try to stick to the established dialectical means for achieving rhetorical objectives which are possibly at odds with the dialectical rationale for a certain discussion rule, and “stretch” the use of these means so much that the fallacious maneuvering is also covered. (Van Eemeren, 2010, p. 140)

Als gevolg van dit doelbewust ‘oprekken’ en uitbuiten van dialectisch geëigende middelen om retorische doelen te bereiken, zullen ontsporingen van strategisch manoeuvreren sterk lijken op redelijke strategische manoeuvres, zodat het in alledaagse discussies lang niet altijd even kristalhelder is waar de grens tussen redelijk versus drogredelijk strategisch manoeuvreren precies getrokken moet worden. Anders uitgedrukt: het onderscheid tussen redelijk en drogredelijk strategisch manoeuvreren is vaak geen simpele zwart-wit kwestie. De verschillende vormen van strategisch manoeuvreren die in de argumentatieve praktijk van alledag onderscheiden kunnen worden, bestrijken dikwijls een continuüm dat loopt van evident drogredelijke tot evident redelijke varianten van strategisch manoeuvreren (Van Eemeren, Garssen, & Meuffels, 2015). Dat geldt ook voor het *argumentum ad consequentiam*. Er blijken verschillende varianten van het *argumentum ad consequentiam* te bestaan die treffende overeenkomsten vertonen met andere wijzen van strategisch manoeuvren waarin een standpunt op

redelijke wijze wordt verdedigd door te wijzen op de consequenties ervan. Zo is er een variant van het *argumentum ad consequentiam* die qua voorkomen lijkt op pragmatische argumentatie en er is ook een variant die veel uiterlijke overeenkomsten vertoont met de redeneervorm *reductio ad absurdum* (Garssen, 2006).

Dit is een voorbeeld van *ad consequentiam*-argumentatie die lijkt op pragmatische argumentatie:

Het onderzoek naar verschillen in intelligentie tussen rassen kan niet kloppen want de resultaten daarvan werken discriminatie van bepaalde groepen in de hand.

In de pragma-dialectische typologie van argumentatieschema's wordt pragmatische argumentatie, een subcategorie van causale argumentatie, als volgt schematisch weergegeven:

1 Standpunt: X moet worden uitgevoerd

1.1 Want: X zal tot het positieve gevolg Y leiden

(1.1') (En: Als X tot positieve gevolgen van het type Y leidt, dan moet X worden uitgevoerd)

(Van Eemeren, 2017, p. 23)

Door twee gemeenschappelijke kenmerken lijkt het *argumentum ad consequentiam* op pragmatische argumentatie: de hypothetische causale claim en de negatieve of positieve waardering die aan de geclaimde gevolgen wordt verbonden. Zowel in pragmatische argumentatie als in het *argumentum ad consequentiam* komt het argument neer op een voorspelling dat wat in het standpunt wordt genoemd een positief of negatief effect zal hebben. Uiteraard zijn er ook verschillen. Terwijl er in pragmatische argumentatie een prescriptief (appellerend) standpunt wordt verdedigd, is het standpunt in het *argumentum ad consequentiam* altijd descriptief van aard. Het is juist de aard van het standpunt die maakt dat pragmatische argumentatie in principe deugdelijk kan zijn terwijl *ad consequentiam* argumentatie qualitate qua ondeugdelijk is. Sommige vormen van *ad consequentiam*-argumentatie kunnen dus als een ontsporing worden gezien van het gebruik van pragmatische argumentatie: de discussiant die zich aan deze drogreden schuldig maakt gaat te ver in het strategisch manoeuvreren en verandert het oorspronkelijke descriptieve standpunt in een prescriptief standpunt.

De tweede variant van *ad consequentiam*-argumentatie lijkt niet op pragmatische argumentatie maar veeleer op een specifieke toepassing van de formeel-logische redeneervorm *modus tollens* die bekend staat als het "bewijs

uit het ongerijmde” ofwel als een redenering *ad absurdum*. Dit is een verzonden voorbeeld van een drogredelijke *ad consequentiam*-argumentatie:

De evolutietheorie kan niet kloppen,
Volgens die theorie zouden we van de apen afstammen
En dat idee is verschrikkelijk.

De volgende argumentatie is een voorbeeld van een niet-drogredelijke toepassing van *ad absurdum*-argumentatie:

De evolutietheorie kan niet kloppen,
Volgens die theorie zouden we van de apen afstammen
En uit genetisch onderzoek blijkt dat dit laatste helemaal niet het geval is.

Wat opvalt als we de *ad consequentiam*- en *ad absurdum*-argumenten vergelijken, is dat het standpunt niet verschilt – dat is in beide gevallen descriptief – maar alleen de laatste premisse. In de niet-drogredelijke *ad absurdum*-vorm is dit een descriptieve uitspraak (de feitelijke ontkenning van de consequens), terwijl de premisse in de *ad consequentiam*-vorm juist evaluatief is.

De beide varianten van het *argumentum ad consequentiam* en hun redelijke tegenhangers zijn weergegeven in Tabel 1.

Tabel 1 Vergelijking van de pragmatische variant van het *argumentum ad consequentiam* (*ad consequentiam* I) en het pragmatische argumentatieschema en van het *argument ad consequentiam* (*ad consequentiam* II) en de *reductio ad absurdum*. Bij zowel pragmatische argumentatie als bij de *reductio ad absurdum* zijn ook negatieve vormen mogelijk.

Pragmatische argumentatie	<i>Ad consequentiam</i> I
Standpunt: X moet worden uitgevoerd	Standpunt: X is waar
want	<i>Want</i>
X leidt tot het positieve gevolg Y	X leidt tot het positieve gevolg Y
en	<i>en</i>
(Als X tot gevolgen van type Y leidt, dan moet X worden uitgevoerd)	(Als X tot gevolgen van type Y leidt, dan is X waar)
<i>Reductio ad absurdum</i>	<i>Ad consequentiam</i> II
Standpunt: X is waar	Standpunt: X is waar
want	<i>want</i>
Als X niet waar is, is Y waar	Als X niet waar is, is Y waar
en	<i>en</i>
Y is niet waar	Y is onwenselijk

Deze twee varianten van het *argumentum ad consequentiam* zijn beide onderzocht in ons grootscheepse onderzoek naar de conventionele deugdelijkheid van de pragma-dialectische discussieregels. De resultaten wijzen uit dat de beide varianten als zeer onredelijk werden beoordeeld (Van Eemeren, Garssen & Meuffels, 2009, pp. 176-179).

Het feit dat de drogreden *ad consequentiam* qua verschijningsvorm lijkt op bepaalde redelijke tegenhangers biedt al een eerste verklaring voor de mogelijke identificatieproblemen die in de praktijk optreden. Door gebruik te maken van specifieke presentatiemiddelen kan de spreker die zich schuldig maakt aan een *ad consequentiam* de drogreden minder herkenbaar maken door deze meer op de redelijke tegenhanger te doen lijken (Garssen, 2016, p. 251). In feite komt dit neer op het creëren van een ambiguïteit: er worden twee interpretatiemogelijkheden geschapen: een redelijke en een onredelijke. Doordat taalgebruikers bij ontstentenis van aanwijzingen dat dit niet terecht is, uitgaan van de presumptie van redelijkheid zal eerder voor de redelijke interpretatie gekozen worden.

Bij de pragmatische *ad consequentiam*-variant wordt gedaan alsof het oorspronkelijke standpunt niet descriptief maar prescriptief-appellerend is. Bij het strategisch manoeuvreren dat er in deze variant op gericht is dat te camoufleren gaat het erom het standpunt op zo'n manier te presenteren dat het als een descriptieve, maar ook als een prescriptieve uitspraak kan worden geïnterpreteerd. In het volgende voorbeeld formuleert Moerman-arts J. Wiese zijn standpunt zo dat beide lezingen mogelijk zijn:

Het is onzorgvuldig uitgedrukt te zeggen dat kanker vooral een kwestie is van onstuitbare celgroei die alleen maar door quasi-heroïsche ingrepen van artsen kan worden gestopt. Sterker nog: dat is zelfs leugenachtig, want deze onbewezen theorie ontnemt de patiënt het vertrouwen in zijn zelf-genezende vermogens, die van binnenuit de kanker kunnen opruimen (Van Eemeren & Snoeck Henkemans, 2011, p. 201).

Door de frasering “het is onzorgvuldig uitgedrukt” kan het standpunt op twee manieren worden uitgelegd: “kanker is niet vooral een kwestie van onstuitbare celgroei” en “het is niet zorgvuldig uitgedrukt om te zeggen dat...” Als we het standpunt in de eerste lezing opvatten, dan levert het argument dat het naar buiten brengen van deze zienswijze tot onwenselijke gevolgen leidt een *ad consequentiam*-drogreden op. Als we het standpunt in prescriptieve zin opvatten, dan past het argument er wel bij. Omdat de spreker daarnaast van “leugenachtig” spreekt, lijkt de descriptieve interpretatie van het standpunt voor te hand te liggen.

Deze interpretatieproblematiek komt uitsluitend voort uit de ambigue formulering van het standpunt, die het lastig maakt de *ad consequentiam*-drogreden meteen bloot te leggen.

Om de pragmatische variant van de *ad consequentiam* moeilijker herkenbaar te maken, is het gezien het voorgaande zaak de aard van het standpunt op een ambigue wijze te formuleren. Dit kan onder meer bereikt worden met behulp van fraseringen zoals *x moet (niet) gezien worden als y* of *x moet (niet) als y worden beschouwd*. In het bovenstaande voorbeeld had dus ook kunnen staan: *kanker moet niet gezien worden als een kwestie van onstuitbare celgroei* (ambigue formulering) in plaats van *kanker is niet een kwestie van onstuitbare celgroei* (niet-ambigue formulering).

We moeten constateren dat de spreker bij het gebruik van deze variant in feite meer drogredenen tegelijk begaat. Met het onduidelijk formuleren van het standpunt overtreedt de spreker de taalgebruiksregel. Als eerder al duidelijk was dat de spreker een descriptief standpunt inneemt en deze nu opeens een prescriptief standpunt lijkt te beargumenteren, dan is er tevens sprake van een overtreding van de standpuntregel, omdat de spreker in feite het oorspronkelijke standpunt aanpast.

Bij de *ad absurdum*-variant van de *ad consequentiam*-drogreden wordt niet zozeer het standpunt gemanipuleerd, maar de aard van het argument. Wordt bij de *ad absurdum*-variant in het argument gezegd dat de consequentie van wat er in het standpunt wordt gezegd onwenselijk is, in de redelijke *ad absurdum*-tegenhanger wordt beweerd dat de consequentie van wat er in het standpunt wordt gesteld onwaar is. Bij het strategisch manoeuvreren kan de spreker dus op zoek gaan naar fraseringen die beide interpretaties (onwaar en onwenselijk) mogelijk maken. In het eerder gebruikte voorbeeld hierna wordt dit bewerkstelligd door gebruik te maken van het bijwoord “absurd”.

De evolutietheorie kan niet waar zijn,
 Als die theorie wel waar was, zouden we van de apen afstammen
 En dat idee is absurd.

“Absurd” zorgt in dit geval voor de beoogde ambiguïteit omdat het zowel “dat idee is niet waar” als “dat is een verschrikkelijk idee” kan betekenen (Garssen, 2016, p. 251)

Kortom, bij het gebruik van beide varianten kan gepoogd worden de *ad consequentiam*-drogreden minder te laten opvallen door twijfel te zaaien over wat de juiste interpretatie is.

Het experiment

In het experiment hebben we ons geconcentreerd op de pragmatische variant van het *argumentum ad consequentiam*; over de resultaten van het onderzoek naar de *ad absurdum*-variant zullen we later rapporteren. In ons experiment werd de volgende hypothese getoetst: *ad consequentiam*-drogredenen worden als minder onredelijk beoordeeld als ze deel uitmaken van argumentatie waarin het oorspronkelijke standpunt als een prescriptief standpunt wordt gepresenteerd. De proefpersonen kregen in totaal 35 discussiefragmenten voorgelegd, waarvan sommige drogredenen bevatten en andere niet. Bij elk fragment moesten ze op een 7-punt Likerttype schaal aangeven hoe (on)redelijk ze de laatste bijdrage aan de discussie vonden. De schaal liep van ‘zeer onredelijk’ (= 1) naar ‘zeer redelijk’ (= 7). We hebben gebruikt gemaakt van een *multiple message design*: van elke type dialoog werden 5 items opgenomen. Een *multiple message design* vergroot de externe en interne validiteit van het onderzoek (van Eemeren, Garssen en Meuffels, 2009, p. 63).

Materiaal

In de laatste bijdragen aan de 35 geconstrueerde discussiefragmenten waren 7 verschillende typen drogredelijke en niet-drogredelijke discussiebijdragen opgenomen, elk vertegenwoordigd door 5 items:

- (1) openlijke *ad consequentiam*-zetten
- (2) vermomde *ad consequentiam* van het pragmatische type-zetten
- (3) deugdelijke zetten met pragmatische argumentatie tab
- (4) deugdelijke zetten met andere dan pragmatische argumentatie
- (5) de *circumstantial*-variant van de *ad hominem*-drogreden
- (6) de *tu quoque*-variant van de *ad hominem*-drogreden
- (7) de *abusive*-variant van de *ad hominem*-drogreden.

De opbouw van deze discussiefragmenten was gefixeerd: alle 35 discussiefragmenten bestonden uit 3 gespreksbeurten, waar telkens 2 discussianten bij betrokken waren. Elk fragment werd voorafgegaan door een korte beschrijving van de context. Uit deze beschrijving blijkt steeds duidelijk welk standpunt oorspronkelijk de aanleiding van de discussie vormt. Bij de *ad consequentiam*-drogredenen is dit standpunt altijd descriptief.

Net als in ons eerdere empirisch onderzoek naar de (on)redelijkheid van drogredenen werden geladen onderwerpen vermeden. In de eerste gespreksbeurt brengt de protagonist een standpunt naar voren, ondersteund

door een argument. Bij de verstopte *ad consequentiam* items is het door de ambigue formulering ook mogelijk het standpunt als een prescriptief standpunt te interpreteren. In de tweede gespreksbeurt maakt de antagonist expliciet of impliciet duidelijk het standpunt niet te aanvaarden. In de derde beurt geeft de protagonist een argument voor het standpunt. Een voorbeeld van een item met een verstopte *ad consequentiam*-drogreden is de volgende:

Sanne en Alex hebben een discussie over de vraag of mannen rationeler zijn dan vrouwen.

Sanne: Rationaliteit en analytisch vermogen kunnen niet worden gezien als mannelijke kenmerken.

Alex: Hoezo niet?

Sanne: Als we het zo bekijken, zouden we mannen onbedoeld een voordeel geven bij sollicitaties en promoties.

Uit de cursief afgedrukte contextbeschrijving blijkt dat de discussanten een descriptief standpunt bediscussiëren. Door de formulering van Sanne in de eerste beurt wordt het standpunt ambigue en zou het ook als een prescriptief standpunt kunnen worden geïnterpreteerd.

Een openlijke *ad consequentiam*-drogreden wordt als volgt tot uitdrukking gebracht:

Lisa en Yvon hebben het over de toekomst van het leven op aarde.

Lisa: Het is aannemelijk dat de zon nooit op zal branden.

Yvon: Hoezo?

Lisa: In dat geval wordt het leven op aarde toch onmogelijk?

In dit laatste item is het standpunt niet ambigue en kan het onmogelijk als prescriptief worden opgevat. Dat maakt de argumentatie die in dit type item wordt gebruikt tot een openlijke *ad consequentiam*.

De toegevoegde items van het type (5) tot en met (7) hebben een dubbele functie. In de eerste plaats fungeren ze als “fillers”, bedoeld zijn om de ware aard van het experiment te verdoezelen: de respondenten mochten niet doorkrijgen dat het ons alleen om hun mening over de *ad consequentiam*-drogredenen begonnen was. In de tweede plaats bieden deze items een controlemogelijkheid van de betrouwbaarheid van het experiment.

Respondenten

Aan het onderzoek deden 31 volwassen vrijwilligers mee, die geselecteerd werden door studenten Taal en Communicatie van de Universiteit van

Amsterdam. De proefpersonen waren allen ouder dan 18 jaar en hadden geen onderwijs genoten over argumentatietheorie. De questionnaire werd schriftelijk afgenomen. Er was een korte schriftelijke instructie; verdere mondelinge uitleg werd niet geboden. Er waren geen aanwijzingen dat de proefpersonen de uitleg niet snapten. De proefleiders kenden de hypothese niet.

Resultaten

De resultaten voor de drie fillers die in Tabel 2 zijn opgenomen, komen qua redelijkheid overeen met de resultaten die we in het project *Conceptions of Reasonableness* bij herhaling hebben gevonden. De *abusive*-drogreden wordt ook nu als de meest onredelijke argumentatieve zet beoordeeld, vervolgens de *circumstantial*-aanval en tot slot de *tu quoque*-variant; opnieuw tenderen deze laatste twee typen drogredenen naar het neutrale middelpunt op de 7-puntsschaal. Kortom, de redelijkheidsscores voor de 3 typen fillers vormen een goede indicatie van de betrouwbaarheid van het experiment. Bovendien zijn beoordeelde onredelijkheid van openlijke *consequentiam*-drogredenen en de beoordeelde redelijkheid van deugdelijke argumentaties geheel in overeenstemming met de bevindingen in het project *Conceptions of Reasonableness*.

Tabel 2 Gemiddelde redelijkheidsscores van de drie filleritems.
Tussen haakjes is de standaarddeviatie weergegeven.

Type argumentatie	Redelijkheidsscore
<i>Ad hominem abusive</i>	2.9 (0.8)
<i>Ad hominem tu quoque</i>	4.1 (1.2)
<i>Ad hominem circumstantial</i>	3.5 (0.7)

Vinden onze respondenten – zoals voorspeld in onze hypothese – de verborgen *ad consequentiam*-zetten inderdaad minder onredelijk dan de openlijke, ‘clear-cut’ gevallen van de *ad consequentiam*-drogreden? De hiervoor relevante data zijn weergegeven in Tabel 3.

Tabel 3 Gemiddelde redelijkheidsscores voor de experimentele items.
Tussen haakjes is de standaarddeviatie weergegeven.

Type argumentatie	Redelijkheidsscore
Pragmatisch redelijk	5.6 (0.9)
Redelijk	5.1 (0.6)
<i>Ad consequentiam</i> openlijk	2.1 (1.2)
<i>Ad consequentiam</i> verstopt	3.1 (1.4)

De gemiddelde redelijkheidsscores voor de vier typen zetten blijken – zoals verwacht – in statistisch significante zin van elkaar te verschillen: $F(2,36) = 35.21$, $p < .01$; $\eta^2 = .48$. Met behulp van een orthogonale *post hoc*-vergelijking contrasteerden we de openlijke *ad consequentiam*-drogreden en de verstopte *ad consequentiam*-drogreden (de voor ons belangrijkste vergelijking). De verstopte *ad consequentiam*-drogreden werd inderdaad minder onredelijk gevonden dan de openlijke *ad consequentiam*-drogreden: $F(1,36) = 17.31$, $p < .01$. In absolute termen vinden de proefpersonen de verstopte *ad consequentiam* nog steeds onredelijk, maar toch 1 punt redelijker dan de *clear cases*. Onze hypothese wordt kortom door de resultaten ondersteund.

Conclusie

De strategische manoeuvre die we in dit experiment hebben onderzocht is complex. Om de onredelijkheid van de pragmatische variant van het *argumentum ad consequentiam* te camoufleren en deze meer op redelijke pragmatische argumentatie te doen lijken, moet er in feite een andere drogreden worden begaan: die van het veranderen van het oorspronkelijke standpunt. Hiermee maken we ons tegelijkertijd aan nog een derde drogreden schuldig: de ambiguïteitsdrogreden. Dit neemt niet weg dat ons onderzoek naar het strategisch manoeuvreren met de *ad consequentiam*-drogreden wel degelijk effecten laat zien. Verpakt als pragmatische argumentatie wordt de *ad consequentiam*-drogreden als minder onredelijk beoordeeld, al wordt ook deze verstopte drogreden nog steeds onredelijk gevonden.

De *ad absurdum*-variant van de *ad consequentiam*-drogreden is in het experiment niet aan de orde gekomen. Het camoufleren van deze variant zal in vervolgonderzoek worden onderzocht.

Noten

1. Zie <https://www.youtube.com/watch?v=AekFGksvuDU>
2. Zie <https://edition.cnn.com/videos/entertainment/2018/05/02/kanye-west-van-lathan-journalist-reacts-nr.cnn>

Bibliografie

- Beevor, A. (2012). *The Second World War*. New York, NY: Little, Brown & Company.
- Eemeren, F.H. van (2010). *Strategic maneuvering in argumentative discourse. Extending the pragma-dialectical theory of argumentation*. Amsterdam-Philadelphia: John Benjamins.
- Eemeren, F.H. van (2017). Argumentative patterns viewed from a pragma-dialectical perspective. In F.H. van Eemeren (Ed.), *Prototypical Argumentative Patterns* (pp.7-29). Amsterdam: John Benjamins.
- Eemeren, F.H. van, Garssen, B., & Meuffels, B. (2009). *Fallacies and Judgments of Reasonableness. Empirical Research Concerning the Pragma-dialectical Discussion Rules*. Dordrecht: Springer.
- Eemeren, F.H. van, Garssen, B., & Meuffels, B. (2012). The disguised 'abusive ad hominem' empirically investigated. Strategic manoeuvring with direct personal attacks. *Thinking and Reasoning*, 18(3), 344-364.
- Eemeren, F.H. van, Garssen, B., & Meuffels, B. (2015). De vermomde ad baculum drogreden empirisch onderzocht. *Tijdschrift voor Taalbeheersing*, 37(1), 79-96.
- Eemeren, F.H. van & Grootendorst, R. (1992). *Argumentation, Communication and Fallacies. A Pragma-dialectical Perspective*. Hillsdale, NJ: Lawrence Erlbaum.
- Eemeren, F.H. van & Grootendorst, R. (2004). *A Systematic Theory of Argumentation. The Pragma-dialectical Approach*. Cambridge: Cambridge University Press.
- Garssen, B. (2016). Face the consequences! Strategic maneuvering with the argumentum ad consequentiam. In F. Paglieri, L. Bonelli, & S. Felletti (Eds.), *The Psychology of Argument. Cognitive Approaches to Argumentation and Persuasion* (pp. 245-253). (Studies in logic: Mathematical logic and foundations 59). London: College Publications.
- Garssen, B. (2006). Beweringen met nare consequenties. Twee varianten van het argumentum ad consequentiam. In B.J. Garssen & A.F. Snoeck Henkemans (Red.), *De redelijkheid zelve. Tien pragma-dialectische opstellen voor Frans van Eemeren* (pp. 107-115). Amsterdam: Rozenberg.

Tekstuele kenmerken van misleidende journalistiek

De zaak-Ramesar

Charlotte Govaert, Luuk Lagerwerf en Céline Klemm
Vrije Universiteit Amsterdam/Monash University, Melbourne

TVT 41 (1): 69–82

DOI: 10.5117/TVT2019.1.005.GOVA

Abstract

Textual characteristics of deceptive journalism: The Ramesar case

Trustworthiness is key in journalism, yet some journalists intentionally deceive their audiences by fabricating sources. In 2014, Dutch journalist Perdiep Ramesar was fired from *Trouw*, an esteemed national newspaper in the Netherlands, after it was discovered that sources in 126 of his articles were ‘non-verifiable’. This poses the question whether untrustworthy news articles may be recognized, for instance by textual characteristics. Using corpus analysis, we searched for systematic differences in source use and presentation comparing Ramesar’s deceptive news articles with two same-sized sets of reliable articles, 1) articles on similar topics from other journalists, and 2) articles with verifiable sources from Ramesar himself. Results indicate that compared to real news sources, fictitious sources are more often secondary definers, who are presented in more stereotypical ways and through more and longer direct quotations.

Keywords: corpus analysis, deceptive journalism, direct quotations, embedded quotations, secondary definers, stereotyping

Inleiding

Een journalist moet bovenal de waarheid vertellen (Kovach & Rosenstiel, 2014); toch zijn er journalisten die de werkelijkheid naar hun hand zetten door verzonnen elementen in hun werk op te nemen. Daarmee ondermijnen

zij het vertrouwen in de journalistiek, terwijl een betrouwbare nieuwsvoorziening van cruciaal belang is voor onze samenleving en, in uiterste instantie, de democratie (Jones Patterson & Urbanski, 2007; Wahl-Jorgensen et al., 2016). De meest voorkomende vorm van misleiding door journalisten is het verzinnen van bronnen, alhoewel journalistieke misleiders ook (tegelijk) feiten verzinnen en/of plagiaat plegen. Dat was althans de uitkomst van een onderzoek van Lasorsa & Dai (2007a) naar het werk van tien leugenachtige Amerikaanse journalisten, onder wie Jayson Blair (*The New York Times*), Patricia Smith (*The Boston Globe*) en Jack Kelley (*USA Today*).

Leugenachtigheid onder journalisten komt niet alleen in Amerika voor. Eind 2014 werd de Nederlandse journalist Perdiep Ramesar ontslagen bij *Trouw* nadat was gebleken dat een aantal van zijn artikelen 'niet-verifieerbare cq onjuiste bronnen' bevatte (Myjer & Smit, 2014, p. 3). Omdat Ramesar medewerking weigerde aan het vervolgonderzoek van rechtsgeleerde Egbert Myjer en de toenmalige Groningse hoogleraar journalistiek Jeroen Smit is onduidelijk of hij bestaande bronnen een andere naam gaf zonder dit te bespreken met een leidinggevende zoals het *Trouw*-handboek (De Berg & Welgraven, 2014) voorschrijft, dan wel een personage smeedde van meerdere bronnen of bronnen volledig uit zijn duim zoog. Hoe dan ook, hij overtrad de meest fundamentele ethische regels van zijn vak die voorschrijven dat een journalist eerlijk, accuraat en transparant hoort te zijn (Society of Professional Journalists, 2014).

Na intern onderzoek zag *Trouw* zich genoodzaakt 126 artikelen van Ramesars hand terug te trekken ('Artikelen met niet-verifieerbare bronnen' z.j.; Myjer & Smit, 2014), een bijzonder pijnlijke kwestie voor een kwaliteitskrant. Dat roept de vraag op of leugenachtige teksten niet eerder te herkennen zijn, bijvoorbeeld aan bepaalde tekstuele kenmerken. Op die manier zouden toekomstige gevallen voorkomen kunnen worden. Om dit te onderzoeken voerden wij een vergelijkende inhoudsanalyse uit, waarbij we Ramesars leugenachtige teksten niet alleen met betrouwbare teksten van zijn eigen hand vergeleken maar ook met artikelen van vergelijkbare lengte over hetzelfde onderwerp van collega-journalisten. Wij wilden hiermee de bestaande kennis over journalistiek bedrog toetsen en uitbreiden aan de hand van een recent, Nederlands geval.

Bronnen in leugenachtige journalistieke teksten

In essentie is journalistiek een 'source-driven process' (Berkowitz, 2009, p. 103). Omdat journalisten op onafhankelijke wijze verslag horen te doen

van de werkelijkheid, beschrijven zij die niet zelf, maar baseren zij zich op bronnen. Journalisten construeren hun teksten dus door middel van attributie (Carlson, 2009). Leugenachtige journalisten hebben de neiging meer (niet minder) bronnen te introduceren dan journalisten die zich aan de regels houden om zo de waarde van hun teksten te vergroten (Lasorsa & Dai, 2007a). We vroegen ons af of dit ook voor Ramesars werk geldt. Onze hypothese luidt dat Ramesar meer bronnen introduceert in zijn leugenachtige dan in zijn waarachtige werk, en ook meer dan collega-journalisten in hun werk (H1).

Leugenachtige journalisten baseren zich niet alleen op meer bronnen, ze raadplegen ook andersoortige bronnen. Atton & Wickenden (2005) maken bijvoorbeeld onderscheid tussen primaire en secundaire duiders, waarbij de eerste groep verwijst naar vooraanstaande figuren of representanten van gerespecteerde instituties in de samenleving terwijl de tweede groep verwijst naar vertegenwoordigers van gemarginaliseerde groepen. Volgens deze systematiek zijn duiders in leugenachtige journalistieke teksten vaker secundair; ze zijn meer divers wat betreft etnische en religieuze afkomst en 'lastig op te sporen' (Lasorsa & Dai 2007a, p. 170). Misleiding met secundaire duiders heeft de nodige voordelen. Waar primaire duiders meer en makkelijker toegang tot de media hebben en zullen klagen als ze opgevoerd worden zonder de journalist in kwestie gesproken te hebben zullen secundaire duiders zoals slachtoffers van mensenhandel, sekswerkers of migranten dit veel minder snel doen. We verwachten dan ook dat Ramesar meer secundaire duiders opvoert in zijn leugenachtige dan in zijn waarachtige werk, ook in verhouding tot zijn collega's (H2).

Secundaire duiders in leugenachtige journalistieke teksten worden vaak op stereotiepe wijze beschreven (Lasorsa & Dai, 2007a; 2007b; Woo, 2003; Gorenfeld, 2004; Shafer, 2002). Stereotypen ontstaan als eigenschappen die doorgaans geassocieerd worden met een sociale groep worden toegeschreven aan een persoon die tot deze groep behoort zonder haar of zijn individualiteit in acht te nemen (Segijn et al, 2014). De verklaring die Lasorsa & Dai geven voor stereotiepe beschrijvingen in leugenachtige journalistieke teksten is dat deze gebaseerd zijn op aannames in plaats van waarnemingen en interacties die daadwerkelijk hebben plaatsgevonden. Leugenachtige stukken zullen bestaande ideeën over bepaalde sociale groepen dus eerder bevestigen dan weerspreken, waardoor lezers en collega's minder snel argwaan krijgen en de leugenachtige journalist minder snel door de mand valt. Onze hypothese is dan ook dat Ramesar bronnen in zijn leugenachtige artikelen in meer stereotiepe bewoordingen beschrijft dan in zijn waarheidsgetrouwe artikelen, ook in vergelijking met collega's

(H3a). Tevens verwachten we dat Ramesar in onbetrouwbare teksten vaker naar de afkomst van bronnen verwijst dan in betrouwbare teksten, en ook vaker dan collega-journalisten (H3b). Uit onderzoek van Van der Linden & Jacobs (2017) blijkt namelijk dat stereotiepe berichtgeving vaak gepaard gaat met verwijzingen naar afkomst.

Tekstuele elementen

We onderzoeken niet alleen wat voor soort bronnen Ramesar introduceert maar ook hoe deze bronnen vervolgens opgevoerd worden. Daarbij richten wij ons in de eerste plaats op *speech presentation*, dat wil zeggen de wijze waarop iemands uitspraken of gedachten weergegeven worden (Semino & Short, 2004). Secundaire duiders spreken vaker dan primaire duiders over hun eigen ervaringen en worden vaker gepresenteerd middels ‘first-person or eyewitness accounts’ (Atton & Wickenden, 2005, p. 349). Dit is een type relaas dat meer dan andere varianten uitnodigt tot een representatie van gedane uitspraken via de directe rede, dat wil zeggen een woordelijke weergave van wat gezegd is. Onwaarachtige nieuwsartikelen maken inderdaad twee keer zoveel gebruik van de directe rede als waarachtige nieuwsartikelen (Dalecki, Lasorsa, & Lewis, 2009). Onze hypothese is dan ook dat Ramesars onbetrouwbare teksten meer directe rede bevatten dan zijn betrouwbare teksten en de teksten van collega-journalisten (H4a). Daarnaast verwachten wij dat de citaten van bronnen die gepresenteerd worden in de directe rede in Ramesars leugenachtige werk langer zijn dan in zijn betrouwbare werk alsmede in dat van andere journalisten (H4b). Eerlijke journalisten moeten namelijk zorgvuldig (gedeelten van) werkelijk gedane uitspraken selecteren om deze in een waarachtig relaas te voegen terwijl oneerlijke journalisten deze restrictie niet ervaren.

Hypothese 4a en 4b worden verder ondersteund door onderzoek van Markowitz en Hancock (2014) dat laat zien dat auteurs die hun lezers bedriegen moeite hebben om de gemiddelde frequenties van bepaalde linguïstische kenmerken binnen het genre te benaderen. Zij concluderen dit op grond van onderzoek naar het werk van Diederik Stapel, de Nederlandse sociaalpsycholoog die meer dan 120 wetenschappelijke artikelen op zijn naam schreef waarvan er 55 gebaseerd bleken op verzonden data. In zijn misleidende werk besteedt Stapel bijvoorbeeld meer woorden aan de gebruikte methode en minder aan het afbakenen van de empirische resultaten – twee belangrijke elementen van het wetenschappelijke discours – dan in zijn eerlijke werk.

Het discours binnen de journalistiek is, net als in wetenschap, gericht op het overtuigen van de lezer van de waarachtigheid van het relaas (Broersma, 2010) maar de journalist gebruikt daarvoor andere methoden dan de wetenschapper. In de journalistiek worden directe citaten bijvoorbeeld geassocieerd met waarachtigheid, ze suggereren immers de aanwezigheid van de verslaggever daar waar de nieuwsfeiten zich afspelen (Vis, Sanders, & Spoor, 2015). Directe citaten kunnen dus ingezet worden om lezers van onbetrouwbare nieuwsberichten (ten onrechte) te overtuigen van de betrouwbaarheid van het artikel. In wetenschappelijke artikelen daarentegen worden personen nauwelijks sprekend opgevoerd. Met andere woorden, de wijze waarop bedrog zich kan openbaren is afhankelijk van conventies binnen een bepaald genre (Markowitz & Hancock, 2016).

In dit verband is ook een ander linguïstisch kenmerk van journalistieke teksten relevant. Het gaat om *embedded quote phenomena* (Q), een klein citaat binnen een langere zin, dat meestal gebruikt wordt om waarachtigheid en een zekere objectiviteit te suggereren (Semino & Short, 2004; Semino, Short, & Culpeper, 1997). Q is dus een tweede citaatvorm die kenmerkend is voor het journalistieke discours en een onwaarachtige journalist wellicht voor problemen stelt bij het bepalen van de juiste, bij het genre horende frequentie. We speculeren dan ook dat Ramesars onbetrouwbare artikelen meer Q bevatten dan zijn betrouwbare artikelen en de artikelen van zijn collega's (H5).

Method

We voerden een kwantitatieve inhoudsanalyse uit waarbij we de leugenachtige artikelen van Ramesar vergeleken met twee andere corpora, te weten een corpus bestaand uit andere artikelen van Ramesar zelf en een corpus bestaand uit artikelen van collega-journalisten. Hierbij dient aangetekend te worden dat de (on)betrouwbaarheid van Ramesars artikelen na zijn ontmaskering is vastgesteld door *Trouw* terwijl de betrouwbaarheid van de artikelen van Ramesars collega's op aannames berust; we deden geen aanvullende tests.

In hun rapport beschrijven Myjers en Smit (2014) hoe is vastgesteld welke teksten van Ramesar betrouwbaar zijn en welke niet. Een deel van de 1011 artikelen die Ramesar schreef tijdens zijn dienstverband bij *Trouw* (korte nieuwsberichten niet meegerekend) produceerde hij met of onder begeleiding van collega's. Omdat de collega's bij wie hierover navraag is

gedaan zeggen dat zij geen enkele aanleiding hadden om te twijfelen aan het bestaan van de genoemde bronnen of de juistheid van hun naam, werden deze artikelen als betrouwbaar aangemerkt. Datzelfde geldt voor artikelen waarin de opgevoerde bronnen uitsluitend bekende politici, woordvoerders of wetenschappers zijn, omdat Myjers en Smit ervan uitgaan dat zij aan de bel zouden trekken als ze ten onrechte als bron zijn opgevoerd. Alle andere artikelen werden aan een nader onderzoek onderworpen. In 126 artikelen bleken bronnen voor te komen die niet traceerbaar waren via Google, Facebook of LinkedIn. Hun namen waren evenmin terug te vinden in de telefoongids of De Nederlandse Familienamenbank van het Meertens Instituut. Hierbij tekenen Myjers en Smit aan dat het uiteraard kan voorkomen dat bestaande individuen onvindbaar zijn via sociale media maar dat vergelijkend onderzoek met werk van Ramesars collega's heeft laten zien dat dit bij hen niet of nauwelijks het geval is. Myjers en Smit vermelden ook dat Ramesar in het begin van 2009 door zijn leidinggevende is aangesproken op zijn brongebruik. Per artikel moesten dit er minstens drie zijn. Dit zou gevolgen kunnen hebben voor het aantal bronnen dat Ramesar opvoert in zijn onbetrouwbare artikelen. Er was echter geen verschil tussen het aantal bronnen in Ramesars onbetrouwbare artikelen van voor en van na februari 2009 ($t < 1$).

Elk van de drie corpora in ons onderzoek omvatte daarom 126 teksten, om te beginnen Ramesars onbetrouwbare teksten (aantal woorden: $M = 770,71$; $SD = 558,25$), die *Trouw* inmiddels uit de archieven heeft verwijderd maar nog wel beschikbaar maakt voor wetenschappelijk onderzoek. Voor elke onbetrouwbare tekst zochten we met behulp van LexisNexis een equivalent van Ramesar zelf, met lengte, onderwerp en publicatiedatum als criteria (waarbij het onderwerp als criterium soms moest worden losgelaten; aantal woorden: $M = 699,12$; $SD = 539,33$). Voor elke onbetrouwbare tekst zochten we in andere Nederlandse nationale dagbladen met behulp van LexisNexis een equivalent wat betreft lengte, onderwerp en publicatiedatum (aantal woorden: $M = 629,89$; $SD = 436,21$; geen significant verschil tussen corpora, $F(2, 379) = 2,37$, $p = 0,10$). Deze dagbladen verschillen niet van *Trouw* voor een belangrijk deel van de onderstaande inhoudsvariabelen (Lagerwerf, Schurink, & Oegema, 2011). Brongebruik was geen selectiecriterium voor de controlecorpora.

Alle variabelen zijn handmatig geteld. Voor de variabelen directe citaten en Q werd ook het aantal woorden binnen het citaat geteld. De analyse-eenheid was het artikel.

Het codeboek werd voor analyse vastgesteld. De variabelen zijn als volgt gedefinieerd.

Aantal bronnen. Een bron is een verwijzing naar een individu of instantie die optreedt als de oorsprong van bepaalde informatie, onafhankelijk van de wijze waarop deze informatie gepresenteerd wordt. Dat kan met name in de vorm van een direct dan wel indirect citaat of zoals samengevat door de journalist. Een voorbeeld van dat laatste is de zin ‘De politie doet momenteel nader onderzoek’, waarbij de politie blijkbaar de bron is van deze informatie zonder dat dit expliciet vermeld wordt. Elke bron is eenmaal per artikel geteld, los van het aantal verwijzingen.

Aantal secundaire duiders. Een primaire duider vertegenwoordigt een gezaghebbende instantie in de samenleving zoals een overheidsinstantie, een bedrijf of een NGO. Soms is de primaire duider een man of vrouw die bekendheid geniet vanwege persoonlijke prestaties (denk aan een sportvrouw of acteur). Dit in tegenstelling tot secundaire duiders, die weliswaar met gezag kunnen spreken over hun persoonlijke ervaringen (Atton & Wickenden, 2005) maar onbekend zijn bij het grote publiek.

Aantal stereotiepe bronnen. Als een bron beschreven is aan de hand van eigenschappen die doorgaans aan een bepaalde sociale groep worden toegeschreven, dan geldt deze als een stereotiepe bron. Voorbeelden van stereotyperende eigenschappen zijn beroep, nationaliteit, afkomst (regio binnen een land of een landsgrenzen overstijgend gebied zoals Midden-Oosten), sociaaleconomische status, gender, etniciteit, religie of seksuele voorkeur.

Verwijzingen naar etniciteit. Dit betreft verwijzingen naar afkomst, dat wil zeggen een land, een regio binnen een land of een regio groter dan een land, en/of ras of etniciteit van de bron.

Directe citaten. Als een bron wordt geciteerd aan de hand van woorden en structuren die blijkbaar letterlijk zo geuit zijn, dan geldt dit als een direct citaat. Directe citaten zijn meestal herkenbaar aan aanhalingstekens en een citaatinbedder (Vliegen, 2014). Een voorbeeld is “‘We doen momenteel nader onderzoek,’ zegt een woordvoerder van de politie.’ Soms ontbreken deze kenmerken echter. Op grond van de context bepaalt de codeur dan of het om een direct citaat gaat. Ook het aantal woorden waaruit het citaat bestaat wordt geteld.

Q (embedded quote). Indien een journalist een bron parafraseert en daarbij een of meerdere woorden tussen aanhalingstekens zet om aan te geven dat deze letterlijk gebruikt zijn, dan wordt dit geteld als Q. Soms worden aanhalingstekens echter gebruikt om ironie of andere, niet-letterlijke vormen van taalgebruik aan te geven. In dat geval bepaalt de codeur op grond van de context of er sprake is van Q. Ook het aantal woorden van elke geïdentificeerde Q wordt geteld.

De intercodeurbetrouwbaarheid is vastgesteld door 13,5% van de artikelen (17 per corpus) door een tweede codeur te laten coderen. Aan de hand van Krippendorff's alpha berekenden we dat de overeenkomst varieerde van 0,79 voor het aantal woorden per Q, via 0,90 voor het aantal bronnen per artikel, 0,98 voor verwijzingen naar etniciteit/ras per artikel tot 0,99 voor het aantal woorden per direct citaat. De gemiddelde betrouwbaarheid tussen codeurs was 0,91.

Resultaten

Omdat de coderingen niet normaal verdeeld waren, werd een serie Kruskal-Wallis-toetsen uitgevoerd om verschillen tussen de corpora te evalueren, met effectgrootte uitgedrukt in een Epsilon-kwadraat (E^2 ; Tomczak & Tomczak, 2014). Door middel van paarsgewijze vergelijkingen met Bonferroni-correctie werden verschillen tussen corpora nader gelokaliseerd. De resultaten staan in Tabel 1.

De eerste Kruskal-Wallis-toets gevolgd door paarsgewijze vergelijkingen (Dunns toets) lieten zien dat Ramesars leugenachtige artikelen zoals verwacht meer bronnen hadden dan zijn betrouwbare artikelen. Ramesars onbetrouwbare werk had echter nauwelijks meer bronnen dan het betrouwbaar geachte werk van zijn collega's. De resultaten ondersteunen H_1 dus slechts ten dele.

Een tweede Kruskal-Wallis-toets gevolgd door paarsgewijze vergelijkingen (Dunns toets) liet zien dat Ramesars leugenachtige artikelen, zoals H_2 al voorspelde, vaker gebaseerd waren op secundaire duiders dan zijn waarachtige artikelen en het werk van collega-journalisten. Een derde Kruskal-Wallistoets gevolgd door paarsgewijze vergelijkingen (Dunns toets) toonde aan dat Ramesars leugenachtige werk zoals voorspeld (H_{3a}) vaker gebaseerd was op stereotiepe bronnen dan zijn waarachtige werk en het werk van andere journalisten. Een aanvullende Kruskal-Wallis-toets liet bovendien zien dat Ramesars onbetrouwbare artikelen significant meer verwijzingen naar etniciteit bevatten dan zijn betrouwbare artikelen (H_{3b}). Wat dit betreft vonden we zoals verwacht geen verschil tussen Ramesars waarheidsgetrouwe werk en werk van collega's.

Op grond van de literatuur namen we aan dat Ramesars onbetrouwbare werk meer (H_{4a}) en langere (H_{4b}) directe citaten bevatte in vergelijking met betrouwbaar werk. Dit bleek inderdaad significant het geval. H_{4a} is daarmee aangenomen. Omdat het aantal directe citaten kan samenhangen met de lengte van het artikel (Dalecki, Lasorsa, & Dai, 2009), berekenden

Tabel 1 Gemiddelden (standaardafwijkingen) van tekstenkenmerken en z-waarden voor paarsgewijze vergelijkingen tussen de corpora van Ramesars misleidende artikelen (N = 126), zijn betrouwbare artikelen (N = 126) en artikelen van andere journalisten (N = 126). H-waarden zijn het resultaat van Kruskal-Wallis-omnibustests (met E^2 als effectgrootte; z-waarden zijn van post hoc paarsgewijze vergelijkingen in Dunns toetsen (Bonferroni-correctie voor p-drempels). De hypothesen zijn genummerd en worden voorafgegaan door 'H'.

	Misleidende journalist – betrouwbare journalist		Misleidende journalist – collega's		Betrouwbare journalist – collega's	
	M (SD)	Resultaat	M (SD)	Resultaat	M (SD)	Resultaat
H1 No. bronnen $H(2) = 8,98^*$ ($E^2 = 0,02$)	3,87 (3,24) – 2,95 (1,94)	$z = 2,90^*$	3,87 (3,24) – 3,68 (2,54)	$z < 1$	2,95 (1,94) – 3,68 (2,54)	$z = 2,4^1$
H2 No. secundaire bronnen $H(2) = 91,26^{***}$ ($E^2 = 0,24$)	2,18 (1,57) – 0,56 (0,84)	$z = 8,91^{***}$	2,18 (1,57) – 1,04 (1,67)	$z = 7,8^{***}$	0,56 (0,84) – 1,04 (1,67)	$z = 1,52$
H3a No. stereotiepe bronnen $H(2) = 91,26^{***}$ ($E^2 = 0,24$)	1,46 (1,54) – 0,31 (0,90)	$z = 8,35^{***}$	1,46 (1,54) – 0,18 (0,70)	$z = 9,64^{***}$	0,31 (0,90) – 0,18 (0,70)	$z < 1$
H3b Verwijzing etniciteit $H(2) = 94,97^{***}$ ($E^2 = 0,25$)	3,88 (4,60) – 0,94 (2,08)	$z = 7,92^{***}$	3,88 (4,60) – 0,96 (2,42)	$z = 8,85^{***}$	0,94 (2,08) – 0,96 (2,42)	$z < 1$
H4a No. directe citaten $H(2) = 33,80^{***}$ ($E^2 = 0,09$)	7,48 (4,79) – 4,27 (4,29)	$z = 5,64^{***}$	7,48 (4,79) – 7,33 (7,45)	$z = 1,57$	4,27 (4,29) – 7,33 (7,45)	$z = 4,08^{***}$

	Misleidende journalist – betrouwbare journalist		Misleidende journalist – collega's		Betrouwbare journalist – collega's	
	M (SD)	Resultaat	M (SD)	Resultaat	M (SD)	Resultaat
H4b % woorden directe citaten H(2) = 56,90*** (F ² = 0,15)	40,73 (22,26) – 24,54 (22,99)	z = 6,02***	40,73 (22,26) – 21,36 (16,27)	z = 6,93***	24,54 (22,99) – 21,36 (16,27)	z < 1
H5 No. Q (emb. quotes) H(2) = 23,01*** (F ² = 0,06)	3,25 (2,87) – 1,96 (2,50)	z = 4,33***	3,25 (2,87) – 2,16 (2,62)	z = 3,95***	1,96 (2,50) – 2,16 (2,62)	z < 1
(H5) % woorden Q H(2) = 18,96*** (F ² = 0,05)	1,69 (2,67) – 1,19 (1,73)	z = 2,06	1,69 (2,67) – 0,97 (1,17)	z = 4,35***	1,19 (1,73) – 0,97 (1,17)	z = 2,25 [†]
†: p < 0,1; *: p < 0,05; **: p < 0,01; ***: p < 0,001						

we ook het aantal woorden per direct citaat in relatie tot het aantal woorden per artikel. Ramesars misleidende artikelen bleken een hoger percentage woorden per direct citaat te bevatten dan zijn waarachtige artikelen en artikelen van collega-journalisten, waarmee ook H4b is aangenomen.

Een volgende Kruskal-Wallis-toets gevolgd door paarsgewijze vergelijkingen (Dunns toets) gaf in overeenstemming met H5 aan dat de leugenachtige Ramesar in vergelijking met de waarheidsgetrouwe Ramesar en zijn collega's hogere frequenties van Q had. Wat betreft het aantal woorden in Q per artikel zagen we echter alleen significant verschil tussen Ramesars misleidende artikelen en het werk van zijn collega's, niet tussen zijn misleidende en zijn waarheidsgetrouwe werk.

Discussie

Journalistieke misleiding wordt zowel door journalisten als door nieuwsgebruikers als een bijzonder ernstige kwestie gezien. Een belangrijke vraag is dan ook of onbetrouwbare artikelen in een vroeg stadium te herkennen zijn, bijvoorbeeld aan de hand van bepaalde tekstuele kenmerken. Om een antwoord te kunnen formuleren deden wij een vergelijkende inhoudsanalyse van Ramesars betrouwbare en onbetrouwbare werk alsmede werk van zijn collega's.

Onze eerste hypothese was dat Ramesar meer bronnen introduceert in zijn leugenachtige dan in zijn waarachtige werk, en ook meer dan collega-journalisten in hun betrouwbaar geachte werk. Inderdaad introduceerde Ramesar minder bronnen in zijn waarachtige dan in zijn onwaarachtige teksten maar niet meer dan betrouwbaar geachte journalisten. Dit betekent onder meer dat Ramesar het gemiddeld aantal bronnen alleen haalde als hij zijn lezers bedroog.

Andere resultaten bevestigden wel onze hypothesen. De onwaarachtige Ramesar introduceerde meer secundaire duiders, die hij vaker op stereotiepe wijze beschreef en vaker onder verwijzing naar hun afkomst. Tevens voerde hij in zijn onbetrouwbare werk bronnen op aan de hand van meer en langere directe citaten. Zijn onwaarachtige werk bevatte eveneens meer Q. Hiermee zijn dus een aantal kenmerken van onbetrouwbare journalistieke teksten geïdentificeerd. Uiteraard zijn er betrouwbare, op goed (secundair) brongebruik gebaseerde journalistieke teksten die gebruik maken van veel en/of lange directe citaten of Q. Deze verteltechnieken kunnen echter bewust of onbewust ingezet worden door journalisten die hun toevlucht nemen tot 'fictie', zoals Lemann (2004) het verwoordt, om hun lezers ervan te overtuigen dat de genoemde bron bestaat en de journalist te woord heeft gestaan.

Lemann (2004) merkte ook al op dat journalistiek bedrog vaak een bepaald soort verhaal betreft, namelijk het levendige achtergrondverhaal met een exotisch tintje. Exotisch verwijst dan naar een wereld die de lezer niet uit eigen waarneming kent. Wij kunnen daaraan toevoegen dat dit samenhangt met de voorkeur van misleidende journalisten voor secundaire duiders, zoals de zaak-Ramesar onderschrijft. Secundaire duiders behoren per definitie tot minderheden, dat wil zeggen dat zij deel uitmaken van een sociale groep waarvan leden vaker dan anderen op stereotiepe wijze beschreven worden. Secundaire duiders worden dus eerder gezien als exponent van een sociale groep dan als individu. Een journalist die zijn bronnen baseert op inbeelding zal makkelijker terugvallen op stereotypen dan een journalist die zich baseert op waarneming van en interactie met een individu (cf Lasorsa & Dai, 2007b; Maass et al, 1995; Wigboldus, Semin, & Spears, 2000). Uit dit fenomeen is het grotere aantal stereotiepe beschrijvingen in onbetrouwbare journalistieke teksten te verklaren.

We onderzochten niet alleen de wijze waarop bronnen in leugenachtige teksten worden beschreven maar ook hoe zij aan het woord gelaten worden. Daarbij keken wij onder meer naar *embedded quotes* (Q). Uit ons onderzoek blijkt dat er wat dit betreft significante verschillen bestaan tussen Ramesars onbetrouwbare werk en dat van zijn collega's, maar niet tussen Ramesars betrouwbare en onbetrouwbare werk. Ramesar gebruikte met andere woorden veel Q, of hij bronnen verzoon of niet. Hier toonde zich de meerwaarde van de onderzoeksopzet. Door Ramesars onbetrouwbare werk niet alleen met zijn eigen betrouwbare werk te vergelijken maar ook met dat van collega's werd het makkelijker om vast te stellen of een tekstueel element een kenmerk was van persoonlijke stijl of van bedrog. Voor de ontwikkeling van het automatiseren van de verificatie van bronnen, iets wat nu anders dan een plagiaatcheck (denk aan Turnitin op universiteiten) nog niet mogelijk is maar in de toekomst wellicht wel, is dit een nuttig inzicht.¹ Hogere frequenties van Q kunnen overigens nog steeds een indicatie van bedrog zijn, bijvoorbeeld bij auteurs voor wie Q geen persoonlijk stijlkenmerk is. Nader onderzoek zou dit moeten uitwijzen.

Tot slot merken wij op dat we van tevoren hadden kunnen bedenken dat er geen significante verschillen in de lengte van Q te vinden zouden zijn. Per definitie is Q namelijk korter dan een zin, terwijl een direct citaat (waarvoor we wel een significant verschil in lengte vonden) meerdere zinnen kan beslaan waardoor de bandbreedte groter is. De frequentie van Q is dus in potentie nog steeds een kenmerk van journalistieke misleiding, maar de lengte van Q vermoedelijk niet.

Voor we aan dit onderzoek begonnen vroegen wij ons af waarom Ramesar zijn lezers (inclusief collega's en leidinggevendenden) zeven jaar lang heeft kunnen misleiden. Achteraf kunnen we zeggen dat Ramesar vermoedelijk eerder door de mand was gevallen als zijn leidinggevendenden het artikel van Lemann (2004) hadden gelezen of het advies van Lasora en Dai (2007a) hadden opgevolgd om meer aandacht te besteden aan journalistieke verhalen met secundaire, lastig te traceren duiders. Tegelijk maakt dit onderzoek inzichtelijk waarom het langer kan duren voor dit type bedrog ontdekt wordt. Secundaire duiders worden vaker opgevoerd in een ooggetuigenverslag, wat het gebruik van directe citaten in de hand werkt. Meer dan indirecte citaten of de vertellende wijs suggereren directe citaten persoonlijk contact met een bron. Ook van stereotiepe beschrijvingen gaat de geruststellende werking van waarachtigheid uit; een relaas dat afwijkt van het bestaande beeld zal op meer argwaan stuiten. Eind- en hoofdredacteuren die een goede antenne willen ontwikkelen voor journalistiek bedrog zouden zich hier meer bewust van moeten zijn.²

Noot

1. In 'Deceptive journalism. Characteristics of untrustworthy news items' (in voorbereiding) gaan wij hier uitvoeriger op in.
2. We danken de studenten die deelnamen aan de Ramesar-bachelorscriptiewerkgroepen in 2015 en 2016.

Bibliografie

- 'Artikelen met niet-verifieerbare bronnen.' (ongedateerd). Zie http://static1.trouw.nl/static/asset/2014/Artikelen_met_niet_verifieerbare_bronnen_Ramesar_2007_2014_7708.pdf
- Atton, C., & Wickenden, E. (2005). Sourcing routines and representation in alternative journalism. A case study approach. *Journalism Studies*, 6(3), 347-359.
- Berg, J. de & Welgraven, C. (2014). *Trouw schrijfboek* (5e ed). Amsterdam: Rainbow.
- Berkowitz, D. (2009). Reporters and their sources. In K. Wahl-Jorgensen & T. Hanitzsch (Eds.), *The Handbook of Journalism Studies* (pp. 102-115). New York, NY: Routledge.
- Broersma, M. (2010). The unbearable limitations of journalism. On press critique and journalism's claim to truth. *The International Communication Gazette*, 72(1), 21-33.
- Carlson, M. (2009). Dueling, dandling or dominating? Journalists and their sources. *Sociology Compass*, 3/4, 526-542.
- Dalecki, L., Lasorsa, D., & Lewis, S. (2009). The news readability problem. *Journalism Practice*, 3(1), 1-12.
- DePaulo, B., Lindsay, J., Malone, B., Muhlenbruck, L., Charlton, K., & Cooper, H. (2003). Cues to deception. *Psychological Bulletin*, 129(1), 74-118.
- Gorenfeld, J. (2004). Blood-thirsty Arabs, vigilante Jews. *salon*. Zie http://www.salon.com/2004/03/23/jack_kelley/

- Hauch, V., Blandón-Gitlin, I., Masip, J., & Sporer, S. (2014). Are computers effective lie detectors? A meta-analysis of linguistic cues to deception. *Personality and Social Psychology Review*, 19(4), 307-342.
- Jones Patterson, M., & Urbanski, S. (2006). What Jayson Blair and Janet Cooke say about the press and the erosion of public trust. *Journalism Studies*, 7(6), 828-850.
- Kovach, B., & Rosenstiel, T. (2014). *The Elements of Journalism*. New York, NY: Random House.
- Lagerwerf, L., Schurink, M., & Oegema, D. (2011). Subjectiviteit als taalgebruik in dagbladen en nieuwswblogs. *Tijdschrift voor Communicatiewetenschap*, 39(3), 107-126.
- Lasorsa, D., & Dai, J. (2007a). Newsroom's normal accident? *Journalism Practice*, 1(2), 159-174.
- Lasorsa, D., & Dai, J. (2007b). When news reporters deceive. The production of stereotypes. *J&MC Quarterly*, 84(2), 281-298.
- Lemann, N. (2004, 15 maart). Blair House. *The New Yorker*. Zie <https://www.newyorker.com/magazine/2004/03/15/blair-house>
- Maass, A., Milesi, A., Zabbini, S., & Stahlberg, D. (1995). Linguistic intergroup bias. Differential expectancies or in-group protection? *Journal of Personality and Social Psychology*, 68, 116-126.
- Markowitz, D., & Hancock, J. (2014). Linguistic traces of a scientific fraud. The case of Diederik Stapel. *PLoS ONE*, 9(8), e105937.
- Markowitz, D., & Hancock, J. (2016). Linguistic obfuscation in fraudulent science. *Journal of Language and Social Psychology*, 35(4), 435-445.
- Myjer, E., & Smit, J. (2014). *Het rapport van de onderzoekscommissie brongebruik Trouw*. Zie http://static3.trouw.nl/static/asset/2014/Onderzoeksrapport_bronnengebruik_Trouw_19122014_7707_7725.pdf
- Segijn, C., Bartholomé, G., Pennekamp, S., & Timmers, M. (2014). De afbeelding van statusverschillen in sekse en etniciteit in Nederlandse non-fictieprogramma's. *Tijdschrift voor Communicatiewetenschap*, 42(3), 305-320.
- Semino, E., & Short, M. (2004). *Corpus Stylistics. Speech, Writing and Thought Presentation in a Corpus of English Writing*. London: Routledge.
- Semino, E., Short, M., & Culpeper, J. (1997). Using a corpus to test a model of speech and thought presentation. *Poetics*, 25(1), 17-43.
- Shafer, J. (2002). Fib Newton. The lesson behind the AP's sacking of reporter Christopher Newton (it's not what you think). *Slate*. Zie <https://slate.com/news-and-politics/2002/10/fib-newton.html>
- Society of Professional Journalists. (2014). *SPJ Code of ethics*. Zie <https://www.spj.org/ethicscode.asp>
- Tomczak, M., & Tomczak, E. (2014). The need to report effect size estimates revisited. An overview of some recommended measures of effect size. *Trends in Sport Sciences*, 21(1), 19-25.
- Van der Linden, M., & Jacobs, L. (2017). The impact of cultural, economic, and safety issues in Flemish television news coverage (2003-13) of North African immigrants on perceptions of intergroup threat. *Ethnic and Racial Studies*, 40(14), 2823-2841.
- Vis, K., Sanders, J., & Spooren, W. (2015). Quoted discourse in Dutch news narratives. In A. Lardinois, S. Levie, H. Hoeken & C. Lüthy (Eds.), *Texts, Transmissions, Receptions. Modern Approaches to Narratives* (pp. 152-172). Leiden: Brill.
- Vliegen, M. (2014). 'Allemaal subjectief', sneerde ze. *Tijdschrift voor Taalbeheersing*, 36(2), 197-224.
- Wahl-Jorgensen, K., Williams, A., Sambrook, R., Harris, J. Garcia-Blanco, I., Dencik, L., Cushion, S., Carter, C., & Allan, S. (2016). The future of journalism. *Journalism Practice*, 10(7), 808-814.
- Wigboldus, D., Semin, R., & Spears, R. (2000). How do we communicate stereotypes? Linguistic bases and inferential consequences. *Journal of Personality and Social Psychology*, 78(1), 5-18.
- Woo, W. (2003). Journalism's "normal accidents." *Nieman Reports*, 57(3), 47-53.

Strategisch manoeuvreren met stijl

Een systematische benadering

Ton van Haften en Maarten van Leeuwen
Universiteit Leiden

TVT 41 (1): 83–96

DOI: 10.5117/TVT2019.1.006.HAAF

Abstract

Strategic maneuvering with style: a systematic approach

In the extended pragma-dialectical approach to argumentation (Van Eemeren, 2010), presentational devices are distinguished as one of the three aspects of strategic maneuvering by arguers in argumentative discourse. Language – or stylistic – devices are an important subset of presentational devices. However, analyses of stylistic devices are often ad hoc: in most cases, a *systematic* analysis of these devices employed by arguers is lacking.

This paper has two (interrelated) aims. Firstly, we aim to show how a systematic analysis of stylistic devices can be integrated in a pragma-dialectical analysis of strategic maneuvering in argumentative discourse. To this end, we make use of methodological insights from linguistic stylistics. Secondly, we argue that the proposed integration of pragma-dialectical argumentation theory and linguistic stylistics is of added value for both disciplines. Our approach and its added value is illustrated by a case study: the pleadings that were delivered in a Dutch civil law case.

Keywords: pragma-dialectical argumentation theory, strategic maneuvering, presentational devices, linguistic stylistics, systematic stylistic analysis

Inleiding¹

Binnen het kader van de uitgebreide pragma-dialectiek (Van Eemeren, 2010) wordt ervan uitgegaan dat taalgebruikers, wanneer zij iemand willen

overtuigen, hun argumentatieve zetten zo kiezen en vormgeven dat zij strategisch manoeuvreren tussen enerzijds de wens om te voldoen aan eisen van redelijkheid en anderzijds de wens om zo effectief mogelijk te zijn. Presentatiemogelijkheden vormen een van de drie aspecten van strategisch manoeuvreren. Talige stijlmiddelen vormen een belangrijk onderdeel van deze presentatiemogelijkheden.² De analyses van de wijze waarop taalgebruikers strategisch manoeuvreren met stijlmiddelen zijn echter veelal ad hoc: een *systematische* analyse van stijl ontbreekt vaak (vgl. Fahnestock, 2009).

Dit artikel heeft twee nauw met elkaar verbonden doelstellingen. Ten eerste willen we laten zien hoe een systematische analyse van stijlmiddelen kan worden geïntegreerd in een pragma-dialectische analyse van argumentatief taalgebruik. Daartoe maken we gebruik van methodische inzichten uit de taalkundige stilistiek (Van Leeuwen, 2015; Stukker & Verhagen, te verschijnen). Ten tweede willen wij betogen dat de door ons voorgestelde integratie van uitgebreide pragma-dialectiek en taalkundige stilistiek voor beide benaderingen waardevol is. Onze doelstellingen werken we uit aan de hand van een juridische casus.

Dit artikel is als volgt opgebouwd. In de volgende paragraaf bespreken we kort een aantal relevante concepten uit de pragma-dialectische argumentatietheorie, waarna we in de daaropvolgende paragraaf onze casus introduceren. Daarna beschrijven we hoe stijlmiddelen waarmee strategisch wordt gemanoeuvreerd systematisch kunnen worden geïdentificeerd, en bespreken we de resultaten van de systematische stijlanalyse van onze casus. In de conclusie komen we terug op de vraag hoe naar ons idee een systematische analyse van stijlmiddelen het best kan worden geïntegreerd in een pragma-dialectische analyse van argumentatief taalgebruik, en wat een kruisbestuiving van de uitgebreide pragma-dialectiek en taalkundige stilistiek beide benaderingen te bieden heeft.

De identificatie en de analyse van de strategische functie van argumentatieve zetten

De uitgebreide pragma-dialectische argumentatietheorie (Van Eemeren, 2010) neemt aan dat een taalgebruiker die deelneemt aan een discussie, strategisch manoeuvreert: hij of zij streeft er in de eigen discussiebijdragen naar om in elke discussiefase een evenwicht te bewaren tussen het doel de discussie te winnen, m.a.w. effectief te zijn, en het doel zich te houden aan standaarden voor redelijkheid. Strategisch manoeuvreren komt in alle argumentatieve zetten in drie verschillende aspecten tot uitdrukking: (a)

de selectie uit het topisch potentieel, dat wil zeggen de keuze die gemaakt wordt uit het reservoir van argumentatieve zinnen waarvan in een bepaalde discussie gebruik kan worden gemaakt; (b) de aanpassing aan de verlangens van het auditorium, dat wil zeggen de afstemming op het referentiekader van degenen die de spreker of schrijver wil overtuigen; en (c) de keuze van de presentatiemogelijkheden, dat wil zeggen het benutten van stilistische en andere uitdrukkingsmiddelen die geschikt zijn om het doel van de spreker of schrijver te dienen (Van Eemeren, 2010, pp. 108-118).

Hoewel de drie aspecten van strategisch manoeuvreren analytisch te onderscheiden zijn, komen ze in elke mondelinge of schriftelijke argumentatieve zin samen en manifesteren ze zich tegelijkertijd.³ In dit artikel richten we ons op een systematische analyse van het derde aspect van strategisch manoeuvreren: de inzet van stilistische keuzes.⁴ Een taalgebruiker die strategisch manoeuvreert, zal o.a. zijn formuleringen bewust of onbewust zo kiezen dat hij binnen de grenzen van de redelijkheid zo effectief mogelijk is. Voor elke fase van de discussie – de confrontatiefase, openingsfase, argumentatiefase en afsluitingsfase (Van Eemeren, 2010, pp. 43-45) – mag worden aangenomen dat de formuleringen van argumentatieve zinnen strategisch zijn gekozen, dat wil zeggen: systematisch zijn afgestemd op de strategische doelen die de spreker of schrijver wil bereiken. Bij het onderzoeken van de vraag hoe er strategisch wordt gemanoeuvreerd met argumentatieve zinnen, is het dan ook belangrijk om stilistische keuzes systematisch in ogenschouw te nemen.

Met het begrip ‘argumentatieve zin’ wordt in de pragma-dialectiek elke relevante zin in een kritische discussie aangeduid; dit begrip is ruimer dan het begrip ‘argument’. Elk argument is een argumentatieve zin, maar niet elke argumentatieve zin is een argument; argumentatieve zinnen komen, anders geformuleerd, ook in andere discussiefasen dan in de argumentatiefase voor. Argumentatieve zinnen in de openingsfase van een discussie zijn bijvoorbeeld voorstellen voor het vaststellen van gedeelde procedurele en materiële uitgangspunten (*common ground*) voor het verdere verloop van de discussie.

Om de strategische functie van een specifieke argumentatieve zin in een bepaalde discussiefase vast te stellen, is het van belang om rekening te houden met o.a. de volgende twee factoren (Van Eemeren, 2010, pp. 163-165).⁵ Ten eerste: omdat aangenomen mag worden dat een taalgebruiker strategisch manoeuvreert om een bepaald resultaat of effect te bereiken, is het van belang om dit beoogde resultaat of effect in de analyse te betrekken. Ten tweede moet de institutionele context ofwel het ‘communicatief activiteitstype’ in de analyse worden verwerkt: argumentatieve zinnen vinden

plaats binnen een specifieke communicatieve praktijk waarin de gang van zaken is afgestemd op de institutionele vereisten van het betreffende communicatieve domein. Deze institutionele vereisten bieden voor het uitvoeren van strategische manoeuvres zowel mogelijkheden als beperkingen (Van Eemeren, 2010, pp. 129-162).

Hoe kunnen inzichten uit de taalkundige stilistiek behulpzaam zijn om binnen een pragma-dialectisch kader *systematisch* te onderzoeken hoe er strategisch wordt gemanoeuvreed met stijlmiddelen? En wat kan een dergelijke geïntegreerde aanpak zowel de uitgebreide pragma-dialectiek en als de taalkundige stilistiek opleveren? Dat zullen we nu illustreren aan de hand van de zaak Van Gelder.

De zaak Van Gelder

Op 7 augustus 2016 kwalificeerde Yuri van Gelder zich in Rio de Janeiro voor het eerst in zijn turnloopbaan voor een Olympische finale. Om dit te vieren, ging hij 's avonds naar het Holland Heineken House. Tegen Van Gelder, die de volgende ochtend weer moest trainen met het turnteam, was gezegd geen alcohol te drinken, en Van Gelder had zijn coach via WhatsApp toegezegd rond middernacht terug te zijn. Na zijn bezoek aan het Heineken House, waar hij een aantal biertjes dronk, ging Van Gelder zijn kwalificatie elders in Rio vieren. Hij keerde in de vroege ochtend terug in het verblijf waar ook andere turners waren gehuisvest en die door Van Gelders thuiskomst werden gestoord. De volgende dag werd Van Gelder om 15.00 uur wakker, waardoor hij de geplande teamtraining had gemist.

Van Gelder werd vanwege zijn gedragingen door NOC*NSF uit het Olympisch team gezet en naar Nederland teruggestuurd; een maatregel die volgens NOC*NSF noodzakelijk was, omdat Van Gelder de normen en waarden van de Olympische ploeg “op grove wijze” had overschreden (Persbericht NOC*NSF, 2016). Terug in Nederland vocht Van Gelder in een kort geding (dat rechtstreeks werd uitgezonden op televisie), de beslissing van NOC*NSF aan, om zo alsnog toegang tot de Olympische finale af te dwingen. Tevergeefs: Van Gelder verloor.

Een 'kort geding' is in pragma-dialectische termen een specifiek communicatief activiteitstype dat binnen het communicatieve domein 'juridische communicatie' behoort tot het genre 'beschikken'.⁶ 'Beschikken' is gericht op het beslechten van een verschil van mening door een bevoegde derde partij in plaats van het oplossen van het meningsverschil door partijen zelf.⁷ In een kort geding daagt een partij een andere partij voor de civiele rechter

om een bepaalde eis ingewilligd te krijgen. De rechter neemt – gehoord de argumenten van beide partijen – een besluit over de vraag of de eis moet worden toegewezen of afgewezen. Dit besluit moet gedragen worden door argumentatie die gebaseerd is op interpretatie van de relevante feiten en de toepasselijkheid van een juridisch beginsel, wettelijke regel of overeenkomst. Specifiek voor het kort geding geldt dat het een heel snelle procedure is, bedoeld om op zeer korte termijn een vaak voorlopig besluit te nemen in urgente zaken, waarbij er weinig procesregels van toepassing zijn.

Voor onze systematische stijlanalyse richten wij ons op de pleidooien die de advocaten van Van Gelder (Cor Hellingman) en van NOC*NSF (Haro Knijff) voor het kort geding hadden voorbereid.⁸ We hebben gekozen voor een analyse van de pleidooien, omdat in dit type betoog meestal alle vier de fasen die idealiter in een discussie kunnen worden onderscheiden ook daadwerkelijk aanwezig zijn, en omdat de strategische doelen van pleidooien meestal eenvoudig onafhankelijk te bepalen zijn op grond van o.a. kennis van het activiteitstype. Zodoende vormen zij goed materiaal om onze benadering te illustreren.

Vanwege ruimtegebrek beperken we ons tot de analyse van argumentatieve zetten in de *openingsfasen* in de pleidooien, waarin voorstellen voor het vaststellen van gedeelde procedurele en materiële uitgangspunten worden gedaan (zie ook de vorige paragraaf).⁹ In veel discussies blijft de openingsfase impliciet, maar in juridische procedures is deze vrijwel altijd een expliciet en belangrijk onderdeel van de discussie, omdat partijen voor de rechter moeten verduidelijken van welke feiten en toepasselijke rechtsregels de rechter bij de beoordeling moet uitgaan. In civiele procedures komt daar bij dat de rechter lijdelijk is: hij moet ervan uitgaan dat hetgeen dat door de ene partij als feit is gepresenteerd en door de andere partij niet expliciet is tegengesproken, als feit vaststaat.¹⁰ Voor de advocaat van een partij die voor de civiele rechter staat, betekent dit o.a. dat hij zijn weergave van de omstandigheden of gebeurtenissen waarover het juridische conflict gaat, verbaal zo moet kleuren dat wordt aangestuurd op een voor zijn cliënt zo gunstig mogelijk besluit door de rechter, terwijl hij tegelijkertijd binnen de (mede door de juridische procedure afgedwongen) grenzen van de redelijkheid blijft.

De systematische identificatie en analyse van de strategische functie van argumentatieve zetten

Methode

Om op systematische wijze stilistische middelen op het spoor te komen in de openingsfasen in de pleidooien van Hellingman en Knijff waarmee

beiden hun argumentatieve zetten strategisch vormgeven, hebben we de volgende methodische stappen gezet.

Eerst hebben we de argumentatieve zetten in de schriftelijke pleidooien van beide advocaten gereconstrueerd, wat resulteerde in een analytisch overzicht van de vier discussiefasen: de confrontatiefase, openingsfase, argumentatiefase en afsluitingsfase. Bij deze reconstructie hebben we ook de specifieke dialectische en retorische doelen van elke fase vastgesteld.

Als gezegd richten we ons in dit artikel op een reeks van stilistische keuzes die in verband kan worden gebracht met het specifieke strategische doel dat elk van de advocaten nastreeft in de openingsfase van een kort geding, namelijk de omstandigheden of gebeurtenissen waarover het juridische geschil gaat, verbaal zo kleuren dat er binnen de grenzen van de (juridische) redelijkheid wordt aangestuurd op een zo gunstige mogelijke uitkomst van het geding voor zijn cliënt. Toegepast op de zaak Van Gelder: verwacht kan worden dat de advocaat van Van Gelder, Hellingman, zal proberen de ernst van Van Gelders gedragingen te *minimaliseren*, terwijl van de advocaat van NOC*NSF, Knijff, verwacht kan worden dat hij juist de ernst van de gedragingen van Van Gelder zal proberen te *maximaliseren*. Deze verwachtingen werden bevestigd bij een eerste lezing van de pleidooien.

Vervolgens hebben we in beide pleidooien systematisch strategische stilistische keuzes geïdentificeerd door een combinatie van bottom-up- en top-downanalyse. Daartoe hebben we gebruikgemaakt van de methode zoals beschreven in Van Leeuwen (2015). Voor die tekstdelen in de pleidooien van Hellingman en Knijff die tot de openingsfase kunnen worden gerekend, hebben we in een eerste analysefase, de bottom-upanalyse, bepaalde stijlmiddelen in kaart gebracht aan de hand van de ‘checklist Nederlandse stijlmiddelen’, die Verhagen (2001) op basis van Leech en Short (1981) voor het Nederlands heeft opgesteld.¹¹ In deze checklist staan tal van linguïstische verschijnselen opgesomd die mogelijk relevant zijn voor de stilistische analyse van een tekst. Deze eerste analysefase is een heuristische fase: de checklist kan de analist helpen om ‘bottom-up’ stilistische verschijnselen op het spoor te komen waarvan het vermoeden bestaat dat die (bewust of onbewust) wel eens strategisch konden zijn gekozen.¹²

Het werken met een checklist vormt dus een belangrijk onderdeel van de hier voorgestelde methodische aanpak: in de heuristische bottom-upanalyse helpt een checklist om de analyse systematischer te maken, doordat de analist als het ware wordt gedwongen om een groot aantal stijlmiddelen in de analyse van de argumentatieve zetten te betrekken. Dit betekent niet dat het werken met een checklist een panacee is: aan het

gebruik ervan en wat het werken ermee oplevert zitten zeker beperkingen (zie voor een bespreking hiervan Van Leeuwen, 2015, pp. 24-26 en pp. 87-88). Zo blijft het identificeren van stijlmiddelen ook bij het gebruik van een checklist deels een intuïtief proces op basis van kennis van een of meer analisten. Zijn er meer analisten bij de bottom-upanalyse betrokken, dan is de analyse natuurlijk ook intersubjectiever. Het gebruik van een checklist heeft voor de stilistische analyse van de verschillende discussiefasen afzonderlijk en van een discussie als geheel echter wel een groot heuristisch voordeel: het per discussiefase systematisch langslopen van een checklist ‘dwingt’ de analist als gezegd als het ware om een groot aantal stijlmiddelen in de analyse te betrekken; dit verkleint de kans dat relevante stijlmiddelen over het hoofd worden gezien (Van Leeuwen, 2015, p. 87).

De uitkomst van de eerste fase in de analyse van argumentatieve zinnen was een lange lijst van stilistische middelen met een vermoedelijk strategisch effect. In de volgende, tweede analysefase hebben we deze grove lijst van eerste stilistische vondsten ‘top-down’ verder ingeperkt. Voor het onderzoek zijn we immers geïnteresseerd in de vraag welke stijlmiddelen strategisch worden ingezet voor de gereconstrueerde doelen die een spreker of schrijver beoogt. Daarom hebben we voor elk van de bottom-up geïdentificeerde stijlmiddelen vastgesteld of aannemelijk kan worden gemaakt dat zij bijdragen aan het verwezenlijken van de gereconstrueerde doelen (en hoe dan precies), door op basis van taalkundige analyse het (veronderstelde) effect van de geïdentificeerde stijlmiddelen vast te stellen.¹³ In het geval van de pleidooien van Hellingman en Knijff hebben we ons daartoe voor elke geïdentificeerde stilistische keuze in de formulering van een argumentatieve zin afgevraagd of die bijdraagt aan het minimaliseren of maximaliseren van de ernst van Van Gelders gedragingen. Alleen stijlmiddelen die aan deze strategische doelen konden worden verbonden, bleven na de top-downanalyse over; het is deze selectie aan stijlmiddelen waarover we in de volgende paragraaf rapporteren.

Resultaten

Welke stilistische keuzes maakte de advocaat van Van Gelder bij het formuleren van zijn voorstellen voor de uitgangspunten voor de discussie om de ernst van Van Gelders gedragingen te minimaliseren? Een eerste relevant stijlmiddel dat tijdens de checklistanalyse naar voren kwam, is het gebruik van verkleinwoorden. Opvallend is dat Hellingman Van Gelders stapavond klein maakt door deze te typeren als een avondje. Dit is niet de enige eufemistische formuleringskeuze in Hellingmans pleidooi: zo wordt Van Gelders nuttigen van alcohol neutraal getypeerd als ‘drankgebruik’ en

als ‘ontspanning’ – wat een positieve connotatie heeft in tegenstelling tot een formuleringskeuze als bijvoorbeeld ‘alcoholmisbruik’. Ook gebruikt Hellingman werkwoorden met een eufemistische connotatie: hij stelt dat Van Gelder in de vroege morgen naar binnen ‘sluipt’, waarmee hij aanstuurt op de gevolgtrekking dat Van Gelder bij zijn terugkeer in het hotel geen geluidsoverlast veroorzaakte. Over de houding van Van Gelders teamgenoten zegt hij: ‘wellicht ergerden zij zich’ – daarmee suggererend dat zij niet boos zouden zijn geweest over diens gedrag; de ernst van Van Gelders gedrag wordt hier nog verder afgezwakt door een eufemistisch werkwoord te combineren met een extensiverend bijwoord. Verder stelt Hellingman dat Van Gelder een training mist omdat hij ‘slaapt’ en omdat hij ‘slaap aan het inhalen is’ – wat positiever klinkt dan bijvoorbeeld ‘dat hij zijn roes ligt uit te slapen’. Bovendien stelt Hellingman dat Van Gelder op enig moment meldt dat hij de geplande training niet ‘zal’ bijwonen, hetgeen suggereert dat Van Gelder daar wel toe in staat was. Met andere woorden, ook de keuze voor het modale werkwoord ‘zullen’ verhult dat Van Gelder, wellicht vanwege een kater, de geplande training niet *kon* bijwonen. De ernst van het missen van de *verplichte* ochtendtraining (die Hellingman typeert als ‘ochtendtraining’ (dus zonder adjectief) of ‘geplande ochtendtraining’) wordt ten slotte ook door strategisch gekozen zinsbouw afgezwakt: ‘Hij [Van Gelder] mist één ochtendtraining, maar rust goed uit.’ Door de gekozen zinsbouw wordt de nadruk op dit uitrusten gelegd. Bij een andere zinsbouw was de focus juist op het missen van de ochtendtraining komen te liggen (‘Hij rust goed uit, maar mist één ochtendtraining’).

Zoals te verwachten valt bij het formuleren van de voorstellen voor de materiële uitgangspunten van de discussie worden in het pleidooi van de advocaat van NOC*NSF tal van formuleringskeuzes gemaakt die de ernst van de gedragingen van Van Gelder juist benadrukken. Een interessante, representatieve passage is de volgende:

Van Gelder heeft herhaalde malen instructies van de turnteamleiding niet opgevolgd. Hij is twee keer spoorloos geweest met alle risico's van dien. Hij heeft zichzelf in een toestand gebracht dat hij een training vanwege overmatig drankgebruik en nachtelijke escapades niet heeft kunnen bijwonen. Hij heeft telkens wisselende verklaringen afgelegd over zijn gedragingen. Hij heeft tegen alle afspraken in alcohol gedronken, terwijl hij nog in competitie was en een kans op een medaille had. Van Gelder heeft een slechte invloed op het team. Zijn teamgenoten hebben hem kwalijk genomen zijn warrige en gedesoriënteerde gedrag na zijn nachtelijke avonturen en in het verlengde daarvan zijn houding om lak te hebben aan de regels van het team (...). (...)

Van Gelder heeft de normen en waarden die gelden binnen TeamNL op flagrante wijze overschreden.

Tal van formuleringskeuzes in deze passage sturen in de richting van de conclusie dat Van Gelders gedrag grensoverschrijdend is geweest. Allereerst valt op dat Knijff de nodige intensiverders (Liebrecht, Hustinx, Van Mulken, & Schellens, 2016) gebruikt om de ernst van de gebeurtenissen te versterken. Zo spreekt hij van ‘nachtelijke escapades’ om de stapavond te karakteriseren (waarmee de suggestie wordt gewekt dat het nogal een drankgelag is geweest), Van Gelder was niet ‘weg’ maar ‘spoorloos’, alcohol nuttigen was tegen ‘alle’ afspraken in, hij had ‘lak’ aan de teamregels, en Knijff stelt dat Van Gelder de normen en waarden van TeamNL op ‘flagrante’ wijze heeft overschreden. Ook de keuze voor een meervoudsvorm is interessant: Knijff heeft niet alleen over ‘escapades’, maar ook over ‘gedragingen’ en ‘avonturen’ die tegen de ‘afspraken’ en ‘regels’ van het team ingingen, en waarover hij ‘wisselende verklaringen’ heeft afgelegd. ‘Instructies’ werden ‘herhaalde malen’ niet opgevolgd, hetgeen ‘risico’s’ tot gevolg had. Ook dit veelvuldig gebruik van de meervoudsvorm benadrukt de ernst van Van Gelders gedrag: Knijff benadrukt door zijn formuleringskeuzes dat er sprake is van verschillende vergrijpen en dat Van Gelder daarmee meerdere regels heeft overtreden. Verder valt het gebruik van de werkwoordstijden op: vrijwel de gehele passage is in de voltooid tegenwoordige tijd geformuleerd (Van Gelder ‘heeft ... niet opgevolgd’, ‘heeft wisselende verklaringen afgelegd’, etc.). In de zin ‘Van Gelder heeft een slechte invloed op het team’ zien we echter een kortstondige overgang naar de *onvoltooid tegenwoordige tijd*. Door deze zin in de onvoltooid tegenwoordige tijd te formuleren, wordt benadrukt dat het incident nadelige consequenties heeft gehad die nog steeds voortduren. Interessant is ten slotte ook dat Knijff stelt dat Van Gelder een training niet heeft ‘kunnen’ bijwonen: met het modale hulpwerkwoord ‘kunnen’ wordt subtiel benadrukt dat Van Gelder niet *in staat was* bij de training te zijn – een suggestie die afwezig zou zijn geweest wanneer Knijff gesteld zou hebben dat Van Gelder ‘een training niet heeft bijgewoond’.

Conclusie

In dit artikel hebben we willen illustreren hoe een systematische analyse van stilistische keuzes kan worden geïntegreerd in een pragma-dialectische analyse van argumentatief taalgebruik, door inzichten uit de (uitgebreide) pragma-dialectiek te combineren met (methodische) inzichten uit de

taalkundige stilistiek. Daaraan verbonden hebben we willen laten zien dat de door ons voorgestelde integratie van (uitgebreide) pragma-dialectiek en taalkundige stilistiek beide benaderingen veel te bieden heeft. Op beide kwesties gaan we hieronder nader in.

Naar ons idee kan een systematische analyse van stilistische keuzes binnen een pragma-dialectische analyse van argumentatief taalgebruik het beste plaatsvinden door de volgende analytische stappen te zetten:

- 1 Een argumentatieve reconstructie maken volgens de pragma-dialectische methode van (een passage binnen) de discussie die het object van analyse en beoordeling vormt, in termen van discussiefasen, typen verschil van mening, argumentatiestructuur, verzwegen argumenten, etc. etc., mede in het licht van het communicatief activiteitstype en het domein waartoe de discussie behoort.¹⁴
- 2 Per discussiefase aan de hand van een checklist¹⁵ een systematische bottom-up- inventarisatie maken van vermoedelijk strategisch gekozen stilistische middelen.
- 3 Deze lijst van systematisch geïdentificeerde stijlmiddelen top-down nader analyseren in termen van hun strategische functie en hun effect, in het licht van de dialectische en retorische doelen van de betreffende discussiefase(n).

Het werken met een checklist vormt een belangrijk onderdeel van de hier voorgestelde aanpak: waar analyses van de wijze waarop taalgebruikers strategisch manoeuvreren met stilistische middelen veelal ad hoc zijn (zie de inleiding), helpt een checklist waarin tal van linguïstische middelen staan opgesomd om de analyse systematischer te maken. Dit betekent als gezegd niet dat het werken met een checklist een panacee is: aan het gebruik ervan en wat het werken ermee oplevert zitten zeker beperkingen (zie onder het kopje 'Methode'). Het gebruik ervan heeft voor de stilistische analyse van de verschillende discussiefasen afzonderlijk en van een discussie als geheel echter wel een groot heuristisch voordeel, en biedt de uitgebreide pragma-dialectiek daarmee een nuttig instrument om analyses van presentationele middelen systematischer te maken. Het per discussiefase systematisch langslopen van een checklist verkleint de kans dat een analist relevante stijlmiddelen over het hoofd ziet (Van Leeuwen, 2015, p. 87). In de Yuri-casus zouden we zonder checklist bijvoorbeeld subtiele, min of meer 'verborgen' grammaticale presentationele middelen als het gebruik van werkwoordstijd of gebruik van modale werkwoorden wellicht niet hebben opgemerkt. Bovendien kunnen met behulp van een checklist niet alleen individuele strategische stilistische keuzes beter worden geïdentificeerd

en op basis daarvan geanalyseerd, maar ook eventuele *discussiestrategieën*: meerdere strategische manoeuvres die gezamenlijk en in samenhang worden ingezet om het resultaat van een bepaalde discussiefase of de discussie als geheel te beïnvloeden (Van Eemeren, 2010, p. 47). Zo hebben we laten zien dat de beide advocaten verschillende strategische manoeuvres inzetten die gezamenlijk en in samenhang moeten bijdragen aan het beoogde resultaat dat de rechter de gedragingen van Van Gelder als niet zo of juist als zeer ernstig zal interpreteren. Met behulp van een checklist wordt de kans dat we een dergelijke strategie identificeren en vervolgens kunnen analyseren groter, omdat het risico dat we relevante strategische formuleringsskeuzes, zoals subtiele grammaticale keuzes, over het hoofd zien, kleiner wordt.¹⁶

De in dit artikel geschetste integratie van de uitgebreide pragma-dialectiek en taalkundige stilistiek vormt naar ons oordeel niet alleen een waardevolle aanvulling voor de argumentatietheorie. Het omgekeerde geldt evenzeer. Een systematische taalkundig-stilistische analyse met hulp van een checklist, levert doorgaans een lange lijst van stilistische observaties op. De analist is echter uiteindelijk niet per se geïnteresseerd in *alle* stilistische observaties, maar in een selectie van stijlmiddelen die in eenzelfde richting 'sturen' en in samenhang een bepaald effect sorteren, zoals ook in de door ons beschreven werkwijze tot uitdrukking komt: eerst voert de analist een systematische bottom-upanalyse uit (met een checklist), met als uitkomst een nog tamelijk ongerichte lijst van stijlobservaties waarin de analist vervolgens nog moet systematiseren en selecteren. Dit gebeurt in de top-downanalyse, waarin de analist zich afvraagt: welke stijlmiddelen sorteren een soortgelijk effect? Welke stijlmiddelen kunnen op zinvolle wijze in samenhang worden besproken? Het beantwoorden van deze vragen is een van de lastigste aspecten van taalkundig-stilistisch onderzoek, zoals we ook merken in ons (MA-) onderwijs.¹⁷

Het grote voordeel van de uitgebreide pragma-dialectiek is dat door een theoretisch en empirisch gefundeerde invulling van onder andere de concepten 'strategisch manoeuvreren' en 'communicatief activiteitstype', een redelijk nauwkeurige vaststelling mogelijk is van het dialectische doel en het retorische doel van elke discussiefase in een concrete discussie of betogende tekst. Daardoor wordt het mogelijk om gefundeerde hypothesen op te stellen over de strategische effecten die een discussiant nastreeft. Dergelijke gemotiveerde hypothesen bieden de analist houvast bij het systematiseren en selecteren van stijlmiddelen die de checklist-analyse heeft opgeleverd: het biedt de analist concrete vragen die hij aan de lijst met bottom-up gevonden stijlmiddelen kan stellen (in de zaak-Van Gelder

bijvoorbeeld: zijn er stijlmiddelen waarmee de beschreven gebeurtenissen worden geminimaliseerd of uitvergroot?), en daarmee belangrijke handvatten om tot een betekenisvolle interpretatie van geïdentificeerde stijlmiddelen te komen.¹⁸ Zo beschouwd is een kruisbestuiving van de uitgebreide pragma-dialectiek en taalkundige stilistiek voor beide benaderingen waardevol bij het analyseren en beoordelen van argumentatief taalgebruik.

Noten

1. De auteurs hebben een gelijkwaardige bijdrage geleverd aan het tot stand komen van dit artikel. We danken twee anonieme reviewers voor hun waardevolle commentaar op een pre-finale versie.
2. Ook beeld- en geluidsmogelijkheden kunnen bij de strategische presentatie van argumentatieve zinnen worden gebruikt.
3. De term 'strategisch manoeuvreren' verwijst zowel naar het uitvoeren van een strategische manoeuvre, dat wil zeggen naar het proces van het maken van keuzes met betrekking tot de drie aspecten van strategisch manoeuvreren, als naar de tekstuele uitkomst van een strategische manoeuvre, dat wil zeggen naar het verbale product dat het resultaat vormt van de verschillende keuzes.
4. Voor het aanduiden van talige presentationele middelen gebruiken wij, in navolging van wat gangbaar is binnen de taalkundige stilistiek, 'stijlmiddelen', 'stilistische keuzes' en 'formuleringskeuzes' als uitwisselbare termen.
5. Volgens Van Eemeren (2010, pp. 163-165) zijn de twee andere factoren die in ogenschouw moeten worden genomen bij het bepalen van de strategische functie van een argumentatieve zet: (a) de mogelijke routes die gevolgd kunnen worden om een bepaald effect te bereiken en (b) de verplichtingen of gebondenheden van de discussiërende partijen die mede de argumentatieve situatie definiëren. Vanwege ruimtegebrek zullen we in onze analyse van de casus niet op deze twee factoren ingaan.
6. Activiteitstypen zijn vaak hiërarchisch geordend. Een 'kort geding' is bijvoorbeeld een communicatieve praktijk binnen het domein van juridische communicatie en een subtype van het activiteitstype 'civielrechtelijk geding', dat weer een subtype is van het activiteitstype 'gerechtelijk geding'. Een activiteitstype kan ook geconstitueerd worden door verschillende andere activiteitstypen. Zo wordt een kort geding geconstitueerd door onder andere de activiteitstypen 'pleidooi' (van elk van de partijen) en 'vonnis' (van de rechter).
7. Een voorbeeld van dit laatste is *mediation*.
8. Wij danken Cor Hellingman en Harro Knijff voor het beschikbaar stellen van dit materiaal.
9. In de argumentatiefase wordt vaak teruggerept op materiële uitgangspunten waarvan in de openingsfase van de discussie expliciet is vastgesteld of impliciet is aangenomen dat de discussiepartners het daarover eens zijn, door deze in te zetten als argumenten. Dit is ook in de onderhavige casus het geval.
10. In die zin is de civiele rechter – in tegenstelling tot de strafrechter – doorgaans dus niet primair uit op waarheidsvinding, maar op de afweging van belangen van partijen.

11. De checklist van Verhagen (2001) is integraal terug te vinden in Van Leeuwen (2015). Zie voor een geüpdatete versie ervan Stukker en Verhagen (te verschijnen).
12. Zie Van Leeuwen (2015, pp. 36-39 en pp. 45-88) en Stukker en Verhagen (te verschijnen) voor een uitgebreidere bespreking van hoe het werken met een checklist in de praktijk van stijlonderzoek eruitziet.
13. Een van de manieren om het (veronderstelde) effect van geïdentificeerde stijlmiddelen aannemelijk te maken, is door te werken met denkbare formuleringsalternatieven. Zie voor een concrete illustratie de volgende paragraaf.
14. Zie o.a. Van Eemeren (2010) en Van Eemeren en Snoeck Henkemans (2016).
15. In dit artikel hebben we gebruikgemaakt van de checklist Nederlandse Stijlmiddelen van Verhagen (2001) – de enige checklist die momenteel voor het Nederlands voorhanden is (zie ook noot 11). Maar er zijn natuurlijk andere checklists denkbaar en voor bijvoorbeeld het Engels ook wel beschikbaar (zie Van Leeuwen, 2015, p. 22 voor een overzicht).
16. De integratie van taalkundig-stilistische analyse in de pragma-dialectiek is naar ons oordeel echter niet alleen van *instrumentele* waarde voor de uitgebreide pragma-dialectische argumentatietheorie, tenminste als je – net als wij – uitgaat van stijlanalyse die is gebaseerd op de cognitieve linguïstiek in combinatie met het uitgangspunt van de (al dan niet radicale) ‘argumentativiteit’ van natuurlijke taal (Ducrot, 2009; Stukker & Verhagen, te verschijnen; Verhagen, 2007). De in die benadering uitgewerkte theorie over argumentatieve semantiek en mentale ruimtes kan naar ons oordeel ook een belangrijke *theoretische* bijdrage leveren aan de verdere gedachtevorming over strategisch manoeuvreren. Het is helaas niet mogelijk om in het bestek van dit artikel hier verder op in te gaan.
17. In taalkundig-stilistische termen betreft het hier de vraag wat precies het ‘macroniveau’ van een tekst vormt voor de stijlanalyse in kwestie (vgl. Van Leeuwen, 2015 en Stukker & Verhagen, te verschijnen). Het macroniveau van een tekst is de indruk die die tekst als geheel maakt op de hoorder of luisteraar. Het macroniveau van een tekst kan verschillende aspecten ervan betreffen, de interpretatie, de overtuigingskracht, de aanvaardbaarheid, de begrijpelijkheid, de aantrekkelijkheid etc. ervan. Het macroniveau van een tekst staat tegenover het microniveau ervan; dat zijn de specifieke stilistische keuzes op klank-, woord-, woordgroeps- en zinsniveau. Taalkundig-stilistische analyse houdt zich o.a. bezig met de vraag op welke wijze keuzes op microniveau bijdragen aan de indrukken op macroniveau.
18. Meer technisch geformuleerd (vgl. noot 17) biedt de uitgebreide pragma-dialectiek de analist zodoende concrete handvatten om tot een gefundeerd ‘macroniveau’ van een tekst voor stijlanalyse te komen.

Bibliografie

- Ducrot, O. (2009). *Slovenian Lectures: Introduction Into Argumentative Semantics*. Ljubljana: Pedagoški inštitut.
- Eemeren, F.H. van (2010). *Strategic Maneuvering in Argumentative Discourse: Extending the Pragma-dialectical Theory of Argumentation*. Amsterdam: John Benjamins.
- Eemeren, F.H. van, & Snoeck Henkemans, A. F. (2016). *Argumentatie: Inleiding in het analyseren en beoordelen van betogen*. Groningen: Noordhoff Uitgevers.

- Fahnestock, J. (2009). Quid pro nobis. Rhetorical stylistics for argument analysis. In F.H. van Eemeren (Red.), *Examining Argumentation in Context. Fifteen Studies on Strategic Maneuvering* (pp. 191-220). Amsterdam: John Benjamins.
- Leech, G., & Short, M. (1981). *Style in Fiction: A Linguistic Introduction to English Fictional Prose*. Harlow: Pearson Longman.
- Leeuwen, M. van (2015). *Stijl en politiek: Een taalkundig-stilistische benadering van Nederlandse parlementaire toespraken*. Utrecht: LOT.
- Liebrecht, C., Hustinx, L., Mulken, M. van, & Schellens, P.J. (2016). Krachtige taal: Een literatuurstudie naar taalintensivering in vier onderzoeksvelden. *Tijdschrift voor Taalbeheersing*, 38, 57-79.
- Persbericht NOC*NSF (2016, 9 augustus). Persbericht Yuri van Gelder. Geraadpleegd op 9 april 2018, <https://www.nocnsf.nl/persbericht-yuri-van-gelder>
- Stukker, N., & Verhagen, A. (te verschijnen). *Stijl, taal en tekst: Stilistiek op taalkundige basis*. Leiden: Leiden University Press.
- Verhagen, A. (2001). *Checklist Nederlandse stijlmiddelen*. Interne publicatie Opleiding Nederlandse Taal en Cultuur, Universiteit Leiden.
- Verhagen, A. (2007). *Constructions of Intersubjectivity: Discourse, Syntax, and Cognition*. Oxford: Oxford University Press.

Probleembesprekingen met samenwerkende kleuters

Frans Hiddink
NHL Stenden Hogeschool

TVT 41 (1): 97–111

DOI: 10.5117/TVT2019.1.007.HIDD

Abstract

Problem conversations with young children during group work

Problem conversations with intervening teachers may enhance children's discourse during group work. The few studies focusing on these conversations, however, are normative and experimental in nature, neglecting the joint and sequential nature of these conversations. Therefore, this paper's aim is to give insight into how teachers and a group of children constitute and continue problem conversations. Detailed analysis informed by Conversation Analysis shows that problems are constructed in three patterns which are distinguished by the teacher's reaction to the problem initiation resulting in three different continuations. It is found that the teacher's actions are decisive for how children may contribute to the developing interaction. In particular, the teacher's reaction to both problem initiations and solution proposals determines whether problem conversations interactions are jointly continued. Theoretical and practical implications are discussed.

Keywords: problem-solving, peer interaction, teacher interventions, group work, conversation analysis

Inleiding¹

De laatste decennia zijn veel pogingen ondernomen om leerlingen actief te laten participeren in interacties in de klas waarin ze gezamenlijk praten en denken. De meeste studies richten zich op interacties tussen leerkrachten en de hele klas (Howe & Abedin, 2013). Divers onderzoek laat namelijk

zien dat dialogische gesprekken waarin kinderen actief participeren, bijdragen aan hun leren en ontwikkeling in tegenstelling tot monologische gesprekken waarin ze testvragen van de leerkracht moeten beantwoorden (Cazden, 2001; Walsweer, 2015).

Anderen richten zich op interacties tussen kinderen wanneer zij in kleine groepjes problemen oplossen. Mercer en collega's demonstreren in diverse studies (zie bijvoorbeeld Mercer, Wegerif, & Dawes, 1999) dat de aard van het taalgebruik van kinderen tijdens het samenwerken van groot belang is. Inmiddels is er dan ook brede consensus dat groepswork kan bijdragen aan cognitieve ontwikkeling, taakuitkomsten en taalontwikkeling, mits kinderen ideeën en standpunten naar voren brengen, deze onderbouwen, verschillen bediscussiëren, problemen oplossen en streven naar overeenstemming (Howe, 2010).

Hoewel enkele effectstudies demonstreren dat het handelen van de leerkracht in interventies tijdens groepswork kan bijdragen aan het verloop van het groepsgeprek, zijn dit soort interventies slechts beperkt onderzocht (Webb, 2009). Chiu (2004) laat bijvoorbeeld zien dat minder inhoudelijke en minder sturende hulp tijdens de interventie bijdraagt aan de manier waarop kinderen hun probleem oplossen direct na de interventie. Aanvullend daarop illustreren andere onderzoekers dat het bevragen van leerlingen *op hun ideeën of inconsistenties* en het uitlokken van *redenen* vóórdat de leerkracht oplossingen aandraagt, zorgt dat kinderen meer gedetailleerd uitleg geven, vaker voortbouwen op elkaars redenties en elkaar meer vragen stellen in het volgende groepswork dan wanneer de leerkracht stuurt en instrueert (zie bijvoorbeeld Dekker & Elshout-Mohr, 2004; Gillies & Boyle, 2006).

In al deze effectstudies worden interventies opgevat als momenten waarop de leerkracht zich bij een groepje leerlingen voegt wanneer die leerkracht of de leerlingen een probleem ervaren. Daarom zullen we in dit artikel spreken over probleembesprekingen. Bovendien zijn deze studies normatief en experimenteel van aard. Zij richten zich vooral op leerkrachtpraktijken die vooraf zijn vastgesteld en die de leerkracht moet uitvoeren in gesprekken met oudere samenwerkende kinderen. Er is daarom nog weinig inzicht in de manieren waarop leerkrachten samen *met* kinderen deze probleembesprekingen vormgeven, zeker niet wanneer het samenwerkingsmomenten tussen jonge kinderen betreft. Dit is opmerkelijk omdat we uit Conversatie Analytische (CA-)studies naar probleembesprekingen in andere institutionele contexten weten dat de manieren waarop aan bepaalde acties in probleembesprekingen vormgegeven wordt, het gesprek(svervolg) beïnvloeden. Zo heeft Stivers (2002) gedemonstreerd dat patiënten op ten

minste twee verschillende manieren hun medische problemen naar voren brengen, die elk een andere druk op de arts leggen om het probleem op te lossen, waardoor ook het gespreksverloop verschilt.

Op vergelijkbare wijze zullen we beschrijven hoe groepjes samenwerkende kleuters met hun leerkracht problemen construeren en hoe het verloop van de probleembesprekingen samenhangt met de manieren waarop probleem geconstrueerd wordt. Maar eerst zullen we kort ingaan op voor dit onderzoek relevante CA-studies naar probleembesprekingen.

Achtergrond

Allereerst zijn probleembesprekingen in *groepen* geanalyseerd. Meier (1997) laat bijvoorbeeld zien dat volwassenen in vergaderingen langzaam een gezamenlijke focus ontwikkelen voordat ze problemen bespreken. Aanvullend demonstreert Huisman (2001) dat de aard van problemen kan verschillen, maar dat gespreksdeelnemers problemen toch op een vergelijkbare wijze construeren, namelijk met een *beschrijving van een stand van zaken* en een (impliciete of expliciete) *negatieve beoordeling* daarvan door dezelfde of een andere spreker. Bovendien blijken probleembesprekingen niet lineair te verlopen van probleemfase, verkenningsfase, naar oplossingsfase, maar blijken die fasen door elkaar te lopen (Huisman, 2001; Meier, 1997).

Een andere relevante studie naar probleembesprekingen in groepen richt zich op jonge kinderen. Het onderzoek van Gosen, Berenst en De Gopper (2015) is, naar wij weten, het enige dat rekening houdt met het gezamenlijke en sequentiële karakter van dit soort interacties in de kleuterklas. Ze hebben laten zien dat problemen tijdens het interactief voorlezen van prentenboeken in klassikale interactie tussen de leerkracht en kleuters op vergelijkbare wijze worden geconstrueerd als in vergaderingen van volwassenen. Ook hier worden problemen geconstrueerd door een *beschrijving* en/of een (impliciet) *negatieve beoordeling* van een stand van zaken, resulterend in een bespreking waarin potentiële oplossingen worden beschreven en beoordeeld door de gespreksdeelnemers.

Daarnaast zijn dyadische probleembesprekingen in de klas bestudeerd. In een studie van Elbers, Hajer, Jonkers, Koole en Prenger (2008) wordt gedemonstreerd dat oudere leerlingen aan het begin meer invloed hebben op het gespreksverloop dan in de latere fasen van zo'n gesprek. Zulke gesprekken blijken leerlingen tijdens het zelfstandig werken te initiëren met een globale probleembeschrijving. Leerkrachten stellen echter nauwelijks

vragen om het probleem te verkennen, zoals te doen gebruikelijk in andere institutionele contexten (Robinson, 2003). In plaats daarvan starten de leerkrachten direct met hun uitleg, waarmee ze het probleem van de leerlingen versmallen. Koole (2012) demonstreert bovendien dat leerkrachten daarmee in feite het probleem oplossen dat ze zelf veronderstellen in plaats van de problemen van de leerlingen. De leerlingen maken overigens zelden bezwaar tegen het handelen van de leerkracht en wanneer ze dat wel doen start de leerkracht een nieuwe uitleg waarbinnen de leerlingen opnieuw alleen kunnen erkennen dat ze iets begrijpen of weten.

De vraag die in dit paper centraal zal staan is nu: *Hoe construeren en bespreken samenwerkende jonge kinderen hun problemen met de leerkracht?*

Data en methodologie

De data komen uit een meerjarig onderzoek waarin we als onderzoekers (Pulles, Hiddink, & Herder, 2014) volgens de methodiek van *Educational Design Research* (EDR; Plomp & Nieveen, 2009) samen met leerkrachten hebben getracht optimale condities te ontwikkelen voor het overleg tijdens het samenwerken in vormen van onderzoekend leren.

De acht participerende onderbouwleerkrachten voerden vijf verschillende projecten volgens de aanpak van Verhalend Ontwerpen (Vos, Dekkers, & Reehorst, 2007) uit in hun klas (groep 1 t/m 3). Daartoe ontvingen ze meerdere globale beschrijvingen van een mogelijke werkwijze voor zo'n project in de klas (Herder, Hiddink, Prenger, & Pulles, 2013), waaruit elke leerkracht een passende keuze kon maken. In elk project zijn minstens drie opnames per klas gemaakt van het groepsoverleg tussen kleuters (2-5 kinderen) terwijl ze problemen van personen in de verhaallijn bespraken, resulterend in een totaal van 73 video-opnames.

Op basis van een eerste verkenning zijn de momenten waarop de leerkracht zich mengde in het groepsoverleg geregistreerd (totaal 275 fragmenten in 73 opnames). Om de onderzoeksvraag te beantwoorden zijn die fragmenten geselecteerd en getranscribeerd waarin een handeling van één van de gespreksdeelnemers functioneerde als oplossing (totaal 115 van de 275), wat we hebben opgevat als suggestie ter ondersteuning voor het bereiken van een doel (zie bv. Robertson, 2017). In de overige gevallen bleken de gesprekken vooral organisatorisch of informerend van aard en werd er noch een probleem noch een oplossing besproken. De geselecteerde probleembesprekingen zijn verder geanalyseerd volgens de methodiek van de

Conversatie Analyse (Mazeland, 2003). Daaruit bleek dat problemen die geïnitieerd worden door leerkrachten anders worden geconstrueerd en besproken dan problemen die geïnitieerd worden door leerlingen (Hiddink & Berenst, 2016). In het kader van dit artikel richten we ons echter alleen op de datacollectie van probleembesprekingen na een initiatie door leerlingen. Daarbij gaat het om 36 (van de 115) fragmenten, afkomstig uit 23 video-opnames.

In onze datacollectie varieerde de setting waarin problemen werden aangekaart. In sommige gevallen stond de leerkracht elders in het lokaal. Soms was zij dan in eerste instantie toehoorder, terwijl in andere gevallen de kinderen haar expliciet vroegen om te komen. In nog weer andere gevallen ging de leerkracht bij het groepje zitten, waarna kinderen het probleem initieerden. Aangezien de projecten verschillend van aard waren en leerkrachten hun eigen keuzes maakten in de projecten die ze wilden uitvoeren, verschilde de aard van de problemen ook, zoals duidelijk zal worden in de bespreking van de fragmenten.

Resultaten

Onze analyse van momenten waarop samenwerkende kleuters hun problemen met leerkrachten bespreken, laat zien dat er drie patronen zijn in de manieren waarop problemen worden geconstrueerd en laat ook zien dat deze patronen elk resulteren in een verschillend vervolg in de probleembespreking. Onder gezamenlijke constructie van het probleem verstaan we het initiëren van het probleem en de eerste reactie van de leerkracht die het probleem als zodanig dan ook erkent. Hoewel de aard van de problemen en de setting waarin het probleem wordt aangekaart tussen de fragmenten in onze collectie verschillen, hebben we niet kunnen vaststellen in welke mate deze verschillen de probleemconstructie beïnvloedt.

Onze analyse maakt duidelijk dat leerlingen een probleem aan de orde stellen door een *negatieve beoordeling* (eventueel met een *beschrijving*) van een stand van zaken. Hoewel die initiatie met verschillende praktijken kan worden vormgegeven en mogelijk gerealiseerd wordt na een elicitatie door de leerkracht, impliceert elke probleemininitiatie dat kinderen een verschil ervaren tussen de huidige en de gewenste stand van zaken. We zullen demonstreren hoe de handeling waarmee de leerkracht reageert op de probleemininitiatie van kinderen het vervolg van de probleembespreking beïnvloedt. We hebben drie verschillende typen leerkrachtreacties gevonden, resulterend in

evenzoveel patronen waarmee problemen geconstrueerd worden. We zullen dat illustreren in de volgende paragraaf. Vervolgens bespreken we de patronen in het vervolg van de probleembespreking na elke probleemconstructie.

Probleemconstructie

In verreweg de meeste gevallen in onze data zien we als eerste patroon waarmee gespreksdeelnemers een probleem construeren, een *probleeminitiatie* van kinderen die direct wordt gevolgd door een *instructie* door de leerkracht. Dit is bijvoorbeeld te zien in fragment 1, waarin kinderen samen een politieagent tekenen.

Fragment 1

1	→ J1	COLi::N, zo heb ik niet meer [genoeg ruimte
2		((kijkt J2 aan))
3	Leerkracht	[((komt
4		aangelopen met stiften))
5	→	met z'n tweeën° kijk. ((legt stiften op
6		tafel en schuift papier naar het midden))

Het probleem wordt geïnitieerd door J1 met een *expliciet* negatieve beoordeling van het handelen van J2 (= Colin) in de vorm van een verwijt (regels 1-2). Hoewel J1 zich richt tot J2, behandelt de leerkracht, die bij een ander tweetal staat, dit verwijt als aan haar gericht (regels 5-6). In haar instructie als reactie laat ze zien dat ook zij het handelen van J2 als problematisch behandelt, namelijk als iets waarvoor een oplossing moet komen. Bovendien geeft ze zelf een oplossing voor het probleem: ze schuift het papier naar het midden van de tafel en voegt de kinderen toe dat ze moeten samenwerken.

Het tweede patroon in de manieren waarop problemen geconstrueerd worden, bestaat uit een *probleeminitiatie* gevolgd door een *opschortingsactie* van de leerkracht. In de meeste gevallen realiseert de leerkracht zo'n actie met een ophelderingsvraag waarin aan het probleem als zodanig wordt gerefereerd. Dit zien we bijvoorbeeld in fragment 2, waarin kinderen het oneens zijn over de ligging van een te ontwerpen robot op papier.

Fragment 2

22		(4.0) [((M en de leerkracht komen terug
23		bij haar groepje met J1 en J2))
24	→ M	[(() - op de
25		kop heeft gezet maar dat is niet <u>waar</u> :
26	→ Leerkracht	wat is het probleem dan?

M is weggelopen naar de leerkracht elders in het lokaal om te klagen over de beschuldiging van J1 aan haar adres. Terwijl ze met de leerkracht terugkomt bij haar groepje rondt ze haar beschrijving af, inclusief een negatief oordeel over de situatie (regels 22-25). De leerkracht reageert met een *ophelderingsvraag* over de situatie, waarmee ze laat zien dat zij door te spreken over 'probleem' de situatie ook als mogelijk problematisch behandelt. Door zo'n actie schort ze de oplossingsbespreking als het ware op, zoals we in de volgende paragraaf, zullen demonstreren. In het voorbeeld daar zien we ook de andere gevonden opschortingsactie, namelijk een *suggestie om nader te overleggen*.

Het derde patroon waarmee gespreksdeelnemers een probleem construeren wordt gevormd door een *probleeminitiatie* van de leerlingen die gevolgd wordt door *uitnodiging(en)* van de leerkracht *om oplossingen* aan te dragen, zoals fragment 3 illustreert. Het voorbeeld demonstreert bovendien dat zo'n probleeminitiatie ook door de leerkracht in een vraag-antwoord sequentie uitgelokt kan worden. Dat aspect van de vormgeving van de probleeminitiatie laten we echter nu verder onbesproken. In dit fragment bespreken ze het resultaat van hun zoektocht naar de verdwenen knuffel 'Raai de Kraai'.

Fragment 3

1	J2	of misschien zit ie in de [boom? ((kijkt J1 aan))]
2		
3	Leerkracht	[((schuift aan))]
4	→	hey, maaruh jullie hebben hem niet hier
5		binnen gevonden? ((kijkt J3 aan))
6	→ J1	[↑nee
7	→ J2	[↑nee
8	→ J3	[↑nee
9	→ M1	[↑nee
10	→ Leerkracht	ok (.) nou, en nu? ((kijkt groepje rond))

Terwijl J2 een hypothese formuleert over de mogelijke verblijfplaats van de verdwenen knuffel (regels 1-2) schuift de leerkracht aan bij het groepje dat zojuist in de school gezocht heeft naar Raai de Kraai, voordat ze een bevestigingsvraag formuleert met een negatief element (regels 3-5) omtrent het gevonden hebben van de kraai. Alle kinderen antwoorden tegelijkertijd dat ze hem niet gevonden hebben (regels 6-9). De leerkracht reageert op de samengestelde probleeminitiatie met een ontvangstbevestiging en een *verzoek* aan de leerlingen om *oplossingen aan te dragen* (regel 10). Daarmee impliceert ze dat de stand van zaken ook dusdanig problematisch acht dat er een oplossing voor gezocht moet worden.

Vervolg van de probleembespreking

Elk van de drie patronen van probleemconstructie blijkt tot een ander vervolg te leiden, zoals we hieronder zullen demonstreren.

Het eerste patroon, waarin de leerkracht met een *instructie* op de probleeminitiatie van de leerlingen reageert, leidt in veel gevallen tot een non-verbale acceptatie van de instructie door leerlingen door de instructie uit te voeren. In enkele andere gevallen zien we echter dat leerlingen een (aanzet tot een) bezwaar realiseren, zoals J2 in regel 9 van fragment 4. In de meeste gevallen is de leerkracht al weggelopen (regel 12) en wordt er niet voortgebouwd op zo'n bezwaar, zoals te zien is in fragment 4, dat het vervolg op fragment 1.

Fragment 4

1	J1	COLi::N, zo heb ik niet meer [genoeg
2		ruimte ((kijkt J2 aan))
3	Leerkracht	[((komt
4		aangelopen met stiften))
5	→	<u>met z'n tweeën</u> ° kijk. ((legt stiften op
6		tafel en schuift papier naar het midden))
7		(.) ((staat op en maakt aanstalten om
8		naar andere leerlingen te lopen))
9	→ J2	[ma- ik ((kijkt naar LK))
10	J1	[een handboei (.) kijk en zo'n sleutelgat
11		((tekent))
12	→ Leerkracht	((loopt weg))

In het tweede patroon van probleemconstructie waarin de leerkracht op de probleeminitiatie reageert met een opschortingsactie leidt zo'n reactie in alle gevallen in onze datacollectie tot een probleemverkenning. Zo'n probleemverkenning wordt gekenmerkt door ophelderingsvragen en antwoorden die veelal uit meerdere sequenties bestaan. Leerlingen geven dan met hun antwoorden meer informatie over de problematische stand van zaken, waarna de leerkracht de verkenning afrondt met een (meer nauwkeurige) beschrijving van de probleemsituatie. Dat wordt gedemonstreerd in fragment 5, waarin kinderen zoeken naar de verdwenen knuffel in de gang.

Het probleem wordt geconstrueerd door een (gezamenlijke) probleeminitiatie (regels 3-5), die wordt gevolgd door een opschortingsactie door de leerkracht, in dit geval in de vorm van een suggestie om het probleem nader te bespreken (regels 8-9). Als ze bij de tafel zijn, start de leerkracht vervolgens de probleemverkenning met een ophelderingsvraag over de locaties waar ze gezocht hebben (regel 15). Hierop volgen enkele vraag-antwoordsequenties

Fragment 5

3	Leerkracht	[was ie boven?
4	M	NEE::
5	Leerkracht	nee[:::?
6	J3?	[misschien helemaal boven
7		(0.6)
8	Leerkracht	nou, ga maar even mee om de tafel zitten (.)
9		dan moeten we het er maar even over hebben
10		
11		((kinderen lopen naar hun tafels; 6 regels
12		weggelaten))
13		
14		((ze zitten weer aan tafel))
15	→ Leerkracht	waar hebben jullie nou allemaal gezocht dan?
16	J1	BUITEN IN HET SPEELLOKAAL IN DE KLAS OP DE
17		GANG BO-VEN
18	→ Leerkracht	ok:: (.) en (.) hij was daar allemaal niet?
19	J2	ook op het verkeersplein
20	→ Leerkracht	ook op het verkeersplein en ik geloof dat er
21		ook jongens waren, die hadden een hut gebouwd
22		(.) hadden jullie daar ook gekeken?
23	J2	ja
24	→ Leerkracht	daar was ie ook niet? (.) nou, en nu?

(regels 15-23), waarin de kinderen met hun antwoorden meer informatie geven over de locaties waar de knuffel niet gevonden was. De leerkracht rondt de verkenning af met een samenvatting waarmee de ernst van de situatie versterkt wordt (regel 24). Zo'n verkenning wordt in alle gevallen afgerond met een soortgelijke probleembeschrijving, waarna de leerkracht de leerlingen uitnodigt om oplossingsvoorstellen te doen. Dat gebeurt ook hier (regel 24: 'en nu?'). Daarmee wordt een oplossingsbespreking gestart, zoals we hieronder zullen illustreren.

Na het derde patroon van probleemconstructie, dat bestaat uit een probleeminitiatie van een kind gevolgd door een uitnodiging van de leerkracht om oplossingsvoorstellen te doen, start meteen de oplossingsbespreking. Zo'n oplossingsbespreking kan dan op twee manieren verlopen, afhankelijk van de leerkrachtreactie op het eerste oplossingsvoorstel van de leerlingen.

De eerste manier waarop de oplossingsbespreking na het derde patroon kan verlopen is in een *information delivery-format*, zoals Koole (2012) dat vond in dyadische probleembesprekingen. Een oplossingsbespreking in deze vorm wordt gerealiseerd door een *evaluatie in combinatie met een instructie* van de leerkracht in reactie op een oplossingsvoorstel. We zien

Fragment 6

60	Leerkracht	hoe kun je nu afspreken wie wat gaat zeggen†
61		(0.6)
62	M1	om[ste beurt p[raten
63	Leerkracht	[nou niet doo[r elkaar heen praten
64	→ M2	[door naar elkaar te luisteren?
65	→ Leerkracht	<u>precies</u> heel goed naar elkaar te luisteren
66		en als je zegt oh iemand vertelt wat en ik wil
67		zo ook wat vertellen, leg dan even je hand op
68		tafel
69		(1.0)
70		dan weten de anderen ook dat je wat wilt
71		zeggen
72		(0.5)
73	→	is dat een idee?=-
74	M2	=ja
75	→ Leerkracht	ja?
76		(0.3)
77	→ Leerkracht	nog een minuutje ik ben heel benieuwd wat
78		jullie bedacht hebben
79		(1.2)
80	→ Leerkracht	ja?
81	M2	ja
82	Leerkracht	ok ((LK staat op en loopt weg) kom ik
83		jullie zo halen

dat in fragment 6, waarin kinderen een probleem hebben met de manier van overleggen met elkaar.

De leerkracht nodigt leerlingen uit om een oplossing voor te stellen als start van de oplossingsbespreking (regel 60). In reactie geeft M1 antwoord, terwijl de leerkracht in overlap haar eerdere uitnodiging herformuleert (regels 62-63). Daarmee gaat ze niet in op het voorstel van M1. In de volgende beurt realiseert M2 een ander oplossingsvoorstel (regel 64). M2's rijzende intonatie laat echter zien dat ze de uitnodiging van de leerkracht als een testvraag behandelt. Uit de daaropvolgende uitgesproken *positieve evaluatie* blijkt dat de leerkracht haar vraag ook als zodanig stelde (regel 65). Na een herhaling van het antwoord, realiseert de leerkracht in het vervolg een instructie (regels 65-73 en verder), waarmee kinderen alleen zeer kort kunnen instemmen (zie bv regels 74 en 81). In dit fragment betreft de leerkrachtreactie een *positieve evaluatie* van een oplossingsvoorstel door de kinderen, gevolgd door een instructie. In onze data zagen we echter ook dat leerkrachten afsluiten met een instructie na een *negatieve evaluatie*.

Een oplossingsbespreking kan dus direct volgen na de probleemconstructie of na een probleemverkenning. Interessant daarbij is dat kinderen in instructies in deze sequentiële positie (dus in post-expansies, na een oplossings-

Fragment 7

24	Leerkracht	daar was ie ook niet? (.) nou, en nu?
25	J1	lokken met uh:: stukjes brood
26	→ Leerkracht	lokken met stukjes brood? (.) heb jij brood?
27	J1	nee
28	Leerkracht	↑oh
29	J3	ik ook niet
30	→ Leerkracht	↑nou het is wel op zich een goed idee (.)
31		want vogels vinden brood wel lekker
32	J1	ja
33	Leerkracht	ja, dat is wel waar (.) [ja
34	J3	[() [brood
35	→ Leerkracht	[we moeten dan maar even- ik kijk
36		straks wel even, want misschien hebben wij
37		ineh het kamertje van de meesters en juffen
38		nog wel brood liggen
39	J3	ja:::
40	Leerkracht	ja, zal ik daar zometeen eens even [kijken?
41	J1	[nee, ik=
42		((steekt vinger omhoog))
43	J2	=of koekjes=
44	J3	of, we kijken straks waar een spoor is
45	→ Leerkracht	een spoor? maar jullie zijn al buiten
46		geweest! heb je al een spoor gezien dan?
47	B3	[nee
48	M1	[nee
49	Leerkracht	oh
50	M1	en ook niet in de zandbak was ie
51	→ Leerkracht	was ie ook niet in de zandbak?
52	J2	we hebben nog niet aan de andere kant van de
53		zandbak gekeken!
54	→ Leerkracht	oh, maar ik denk dat jullie allemaal rondjes
55		hebben gerend (.) of niet?
56	J3	nee
57	J1	nee
58	→ Leerkracht	daar hebben jullie niet gekeken?
59	J1	() naar buiten
60	J2	want we hebben niet buiten ()

voorstel door leerlingen) alleen instemmings-tokens realiseren maar geen bezwaren, zoals wel het geval kan zijn bij de leerkrachtinstructies direct na een probleeminitiatie, zoals hierboven beschreven bij fragment 4.

Een tweede manier waarop de oplossingsbespreking na de probleemconstructie volgens patroon 3 kan verlopen, vertoont kenmerken van een 'discussiekader' in klassikale interactie (Walsweer, 2015). In zo'n oplossingsbespreking kunnen kinderen kritisch participeren, waarbij (ook) kinderen voorstellen evalueren, alternatieven aandragen en onderbouwen met argumenten.

Zo zien we -in fragment 7- hoe diverse leerkrachtpraktijken zoals echte (= *informatiezoekende*) *vragen*, *tentatieve evaluaties* als reactie op een oplossingsvoorstel en een aankondiging (voorstel) de bijdragen van leerlingen bevorderen. In fragment 7 bespreken de kinderen met de leerkracht hoe ze kunnen zorgen dat Raai de Kraai, een verdwenen knuffel, weer kan terugkeren.

In dit voorbeeld worden meerdere opties om het probleem van de verdwenen kraai op te lossen tegen het licht gehouden en soms kritisch onderbouwd of afgewezen, vooral door de kinderen, maar vaak wel naar aanleiding van specifieke acties van de leerkracht. In zo'n kader realiseert de leerkracht *informatiezoekende vragen* (*regels 26, 46, 51 en 58*) en *tentatieve evaluaties* (*regels 30-31*) in reactie op de voorstellen van kinderen. Bovendien realiseert de leerkracht *voorstellen* (*regels 35-38 en 40*) en *bezwaren* (*regels 45-46 en 54-55*). Dergelijke leerkrachtacties lijken nadrukkelijk bij te dragen aan een constellatie waarin kinderen actief en kritisch participeren.

Conclusie en discussie

In deze studie onderzochten we manieren waarop groepjes samenwerkende kleuters samen met leerkrachten hun probleem construeren en hoe deze besprekingen vervolgens verlopen.

Zoals we hebben geïllustreerd, vonden we drie patronen waarmee de gespreksdeelnemers problemen construeren. Hoewel de praktijken waarmee de negatieve beoordeling in de probleeminitiatie wordt gerealiseerd, kunnen verschillen, zijn de patronen met name te onderscheiden door de reactie van de leerkracht op de probleeminitiatie waarmee het probleem geconstrueerd wordt. We hebben drie verschillende reactiewijzen onderscheiden, te weten een instructie, een opschortingsactie en een uitnodiging om oplossingen aan te dragen. Die reactiewijzen karakteriseren dus ook drie probleemconstructie-patronen. Het eerste patroon bestaat uit een probleeminitiatie, direct gevolgd door een instructie. In de instructie in deze sequentiële positie zien we dat kinderen soms (aanzetten tot)

bezwaren realiseren, maar de leerkracht gaat daar niet op in. Van een gespreksvervolg is dan slechts heel beperkt sprake. Het tweede patroon bestaat uit een probleeminitiatie, gevolgd door een opschortingsactie. Dit patroon leidt tot een probleemverkenning bestaande uit drie elementen voordat de gesprekdeelnemers de oplossingsbespreking realiseren. Het derde patroon bestaat uit een probleeminitiatie gevolgd door een uitnodiging voor een oplossingsvoorstel. Dit patroon leidt tot een probleemoplossingsbespreking. Zoals we hebben gedemonstreerd, kan zo'n oplossingsbespreking (eventueel voorafgegaan door een probleemverkenning) op twee manieren verlopen. De oplossingsbespreking kan het karakter krijgen van een evaluatie en een instructie als de leerkrachtreactie op een oplossingsvoorstel van leerlingen resulteert in een *information-delivery format* waarin de deelname van de leerlingen slechts beperkt is tot instemmen met de leerkracht. De oplossingsbespreking kan echter ook het karakter krijgen van een discussie waarin kinderen kritisch participeren met evaluaties, argumenten en alternatieve voorstellen en leerkrachten vooral informatiezoekende vragen, tentatieve evaluaties, voorstellen, suggesties en bezwaren realiseren.

Zoals we hebben gedemonstreerd, hangt het vervolg van de probleembespreking vooral af van de leerkrachtreactie op de probleeminitiatie van de leerlingen. Het lijkt erop dat de vormgeving van de probleeminitiatie geen invloed heeft op de leerkrachtreactie. De kinderen realiseren namelijk steeds op een vergelijkbare wijze problemen (eventueel na een elicitatie door de leerkracht), ondanks de grote variatie in problemen in onze datacollectie. Het centrale kenmerk van die initiaties is dat er een stand van zaken wordt beschreven met een (impliciete) negatieve beoordeling, wat vergelijkbaar is met de manieren waarop gespreksdeelnemers in andere groepssituaties problemen initiëren (Gosen et al., 2015; Huisman, 2001).

Onze bevindingen demonstreren ook dat er behalve een directe instructie na de probleeminitiatie door kinderen, zoals Elbers et al. (2008) vonden, ook nog alternatieve patronen om problemen te construeren, te vinden zijn. Deze leiden bovendien tot een ander verloop van het gesprek, waarin jonge kinderen meer kunnen participeren. In de meeste gevallen in onze data reageren leerkrachten echter wel direct met een instructie, waarin leerlingen soms bezwaar aantekenen. Omdat zulke bezwaren de vraag oproepen wiens probleem er nu eigenlijk opgelost wordt, dat van de leerlingen of dat van de leerkracht (Koole, 2012), lijkt het erg aannemelijk dat leerkrachten in onze data met een instructie in die sequentiële positie een oplossing aandragen voor door hen veronderstelde problemen.

Probleembesprekingen tijdens groepswork hebben dus in principe de potentie om kinderen actief en kritisch te laten participeren. Wanneer leerlingen daarin hun eigen ideeën kunnen formuleren, onderbouwen en confronteren met ideeën van anderen, kan dat vervolgens bijdragen aan het onderling overleg (Chiu, 2004). Zoals we hebben laten zien, hangt het wel of niet tot stand komen van zo'n situatie sterk samen met de manier waarop leerkrachten reageren op probleeminitiaties en vervolgens op oplossingsvoorstellen van leerlingen. Vervolgonderzoek zou daarom ook gericht moeten worden op de vraag hoe we leerkrachten zouden kunnen oriënteren op de effecten van hun reactiewijzen. Omdat probleembesprekingen tijdens onderzoekend leren in kleine groepjes heel divers van aard kunnen zijn, menen we dat een praktijkgericht onderzoek interessante mogelijkheden biedt. Walsweer (2015) heeft bijvoorbeeld geïllustreerd hoe onderzoekers kunnen aansluiten bij het perspectief van leerkrachten en met hen praktijken in de eigen klassikale interacties kunnen onderzoeken volgens de methodiek van EDR (Plomp & Nieveen, 2009).

Uit ons onderzoek blijkt ten slotte dat een probleemverkenning niet noodzakelijkerwijs plaats hoeft te vinden in probleembesprekingen tijdens het groepswork bij jonge kinderen. Dit staat in contrast met de normatieve ideeën over dit soort gesprekken, waarin wordt gesuggereerd dat leerkrachten eerst het probleem *moeten* verkennen, voordat er oplossingen worden besproken (Van de Pol, Volman, & Beishuizen, 2012). Wellicht zou de *aard van het probleem* een verschillend appel kunnen doen op een nadere verkenning, al hebben we dat niet kunnen vaststellen. In ieder geval roept onze bevinding op tot terughoudendheid in het voorschrijven van gewenste kenmerken van probleembesprekingen zonder empirische basis van beschrijvend onderzoek.

Noot

1. Ik dank Jan Berenst voor zijn nauwgezette commentaren op een eerdere versie van dit artikel.

Bibliografie

- Cazden, C. B. (2001). *Classroom Discourse: The Language of Teaching and Learning*. Portsmouth, NH: Heinemann.
- Chiu, M. M. (2004). Adapting teacher interventions to student needs during cooperative learning: How to improve student problem solving and time on-task. *American Educational Research Journal*, 41(2), 365-399.

- Dekker, R., & Elshout-Mohr, M. (2004). Teacher interventions aimed at mathematical level raising during collaborative learning. *Educational Studies in Mathematics*, 56(1), 39-65.
- Elbers, E. P. J. M., Hajer, M., Jonkers, M. D. J., Koole, A. J., & Prenger, J. (2008). Instructional dialogues: Participation in dyadic interactions in multicultural classrooms. In J. Deen, M. Hajer, & A. J. Koole (Eds.), *Interaction in Two Multicultural Mathematics Classrooms. Mechanisms of Inclusion and Exclusion* (pp. 139-170). Amsterdam: Aksant.
- Gillies, R., & Boyle, M. (2006). Ten australian elementary teachers' discourse and reported pedagogical practices during cooperative learning. *The Elementary School Journal*, 106(5), 429-452.
- Gosen, M. N., Berenst, J., & de Gloppe, K. (2015). Problem-solving during shared reading at kindergarten. *Classroom Discourse*, 6(3), 175-197.
- Herder, A., Hiddink, F., Prenger, J., & Pulles, M. (2013). *Samen onderzoeken. Drie onderzoeksprojecten voor groep 1 t/m 8*. Leeuwarden: NHL, Lectoraat Taalgebruik & Leren.
- Hiddink, F., & Berenst, J. (2016). Probleemoplossende gesprekken. Van wie is het probleem? *De wereld van het jonge kind*, 44(4), 23-25.
- Howe, C. (2010). Peer dialogue and cognitive development: A two-way relationship? In K. Littleton, & C. Howe (Eds.), *Educational Dialogues. Understanding and Promoting Productive Interaction* (pp. 32-47). Abingdon: Routledge.
- Howe, C., & Abedin, M. (2013). Classroom dialogue: A systematic review across four decades of research. *Cambridge Journal of Education*, 43(3), 325-356.
- Huisman, M. (2001). Decision-making in meetings as talk-in-interaction. *International Studies of Management & Organization*, 31(3), 69-90.
- Koole, T. (2012). The epistemics of student problems: Explaining mathematics in a multi-lingual class. *Journal of Pragmatics*, 44(13), 1902-1916.
- Mazeland, H. (2003). *Inleiding in de conversatieanalyse*. Bussum: Coutinho.
- Meier, C. (1997). *Arbeitsbesprechungen: Interaktionsstruktur, interaktionsdynamik und konsequenzen einer sozialen form*. Opladen: Westdeutscher Verlag.
- Mercer, N., Wegerif, R., & Dawes, L. (1999). Children's talk and the development of reasoning in the classroom. *British Educational Research Journal*, 25(1), 95-111.
- Plomp, T., & Nieveen, N. (Eds.). (2009). *An Introduction to Educational Design Research*. Enschede: SLO.
- Pulles, M., Hiddink, F., & Herder, A. (2014). Taalontwikkeling door onderzoekend leren binnen thema's. *MeerTaal*, 2(1), 14-17.
- Robertson, S. I. (2017). *Problem Solving. Perspectives from Cognition and Neuroscience*. New York, NY: Routledge.
- Robinson, J. D. (2003). An interactional structure of medical activities during acute visits and its implications for patients' participation. *Health Communication*, 15(1), 27-59.
- Stivers, T. (2002). Presenting the problem in pediatric encounters: 'Symptoms only' versus 'candidate diagnoses' presentations. *Health Communication*, 14(3), 299-338.
- Van de Pol, J., Volman, M., & Beishuizen, J. (2012). Promoting teacher scaffolding in small-group work: A contingency perspective. *Teaching and Teacher Education*, 28(2), 193-205.
- Vos, E. J., Dekkers, P., & Reehorst, E. (2007). *Verhalend ontwerpen: Een draaiboek*. Groningen: Wolters Noordhoff.
- Walsweer, A. (2015). *Ruimte voor leren: Een etnografisch onderzoek naar het verloop van een interventie gericht op versterking van het taalgebruik in een knowledge building environment op kleine friese basisscholen*. Groningen: Rijksuniversiteit Groningen.
- Webb, N. M. (2009). The teacher's role in promoting collaborative dialogue in the classroom. *British Journal of Educational Psychology*, 79, 1-28.

Wat leert onderzoek naar overtuigende teksten over het ontwerpen van overtuigender teksten?

Een overzicht van meta-analytische studies

Hans Hoeken
Universiteit Utrecht

TVT 41 (1): 113–136

DOI: 10.5117/TVT2019.1.008.HOEK

Abstract

What are the lessons of research on persuasive documents for the design of more persuasive documents? A review of meta-analytic studies

Persuasive communication is designed to elicit certain thoughts and feelings in order to change the audience's attitudes and intentions, with the ultimate goal of influencing behavior. This paper addresses the question what meta-analyses in the field of persuasion can tell us about the effects that persuasive communication can have on the audience's behavior and on the steps leading up to that behavior. A review of these meta-analyses shows that changing intentions does not automatically lead to similar sized behavioral effects, that communicative interventions only have small to medium effects, that these effects can be obtained through ethos, pathos, and logos, and that these effects are to some extent susceptible to message design characteristics. The (small to medium) size of the effects of communicative interventions has implications for the design of research on persuasive communication as well as for the advice for practitioners.

Keywords: persuasion, message design, meta-analyses, communication strategies

Inleiding¹

In onderzoek naar persuasieve communicatie gaat het veelal om de vraag of verschillen in inhoud, structuur, stijl en vormgeving van een boodschap

tot verschillen in overtuigingskracht leiden. Idealiter leidt dit onderzoek tot een beter inzicht in de wijze waarop boodschapkenmerken het overtuigingsproces sturen en de uitkomst ervan bepalen. Op basis van dit inzicht is het mogelijk om empirisch gefundeerde adviezen te geven aan voorlichters, politici, fondswervers van goede doelen en reclamemakers die het gedrag van hun doelgroep willen beïnvloeden. Het onderzoek zou moeten leiden tot accuratere voorspellingen over welke boodschappen erin zullen slagen om gezondheids-, stem-, geef- en consumentengedrag te beïnvloeden.

O'Keefe (2015) wijst op het nut van meta-analyses voor de onderbouwing van gefundeerde communicatie-adviezen. In de afgelopen jaren zijn er veel relevante meta-analyses verschenen. Op basis daarvan probeer ik de volgende vragen te beantwoorden.

- Communicatie richt zich meestal op de beïnvloeding van gedragsdeterminanten zoals intenties, attitudes, normen en eigen-effectiviteit. In hoeverre leiden veranderingen in dergelijke determinanten tot corresponderende veranderingen in gedrag?
- Veel communicatie-interventies zijn uiteindelijk gericht op gedragsverandering. In hoeverre zijn communicatie-interventies succesvol?
- Van oudsher worden er drie overtuigingsstrategieën onderscheiden: het inzetten van bronkenmerken (ethos), het inspelen op gevoelens (pathos) en het geven van argumenten (logos). In hoeverre zijn deze strategieën effectief in het beïnvloeden van de attitudes en intenties van het publiek?
- Er vindt veel onderzoek plaats naar specifieke boodschapkenmerken zoals het expliciteren van conclusies of het gebruik van metaforen. In hoeverre leidt toepassing van deze boodschapkenmerken tot overtuigender boodschappen?

Voordat ik de bovenstaande vragen bespreek, geef ik eerst een korte beschrijving van wat meta-analyses zijn en wat hun belang is voor onderzoek naar persuasieve communicatie.

De meta-analyse als methode

Veel onderzoek naar persuasieve communicatie wordt gekenmerkt door de volgende twee kenmerken: er worden maar één of twee boodschappen onderzocht én de focus ligt op de significantie van het verschil in plaats van op de grootte ervan. Zo onderzocht Hoeken (2001) de relatieve overtuigingskracht van anekdotische, statistische en causale evidentie in een

krantenbericht over de bouw van een cultureel centrum in de Achterhoek. Hij vond een significant verschil: de claim dat dit een zinvol initiatief is, werd in sterkere mate geaccepteerd als die werd ondersteund door statistische evidentie dan door anekdotische of causale evidentie.

Het eerste probleem is dat dit resultaat niets zegt over hoe het verschil in evidentietype zou uitpakken in een discussie over een ander thema via een ander medium voor een andere doelgroep. Is statistische evidentie ook overtuigender dan anekdotische of causale evidentie als die wordt gebruikt ter onderbouwing van standpunten zoals dat onveilige seks tot geslachtsziekten kan leiden of dat orgaandonatie wenselijk is? In verschillende bijdragen is besproken waarom je de resultaten voor één boodschap niet kunt generaliseren naar andere boodschappen (zie, bijvoorbeeld Eerland & Van den Bergh, 2016; Jackson & Jacobs, 1983; Meuffels & Van den Bergh, 1996).

Het tweede probleem is de focus op significantie. Als de kans dat het gevonden verschil op toeval berust kleiner is dan .05, dan concludeert de onderzoeker dat de manipulatie effect heeft gehad op de overtuigingskracht. Is de *p*-waarde hoger, dan luidt de conclusie dat er géén effect is opgetreden (ongeacht of die *p*-waarde .07 of .87 is). Voor de theorievorming over persuasieve communicatie kan de *significantie* van een effect relevant zijn; voor een gefundeerd praktijkadvies is de *grootte* van effect belangrijker: leidt de aan- of afwezigheid van het boodschapkenmerk tot een fors, gemiddeld of een klein verschil in overtuigingskracht? Voor de praktijk is het boodschapkenmerk met een fors effect interessanter dan het kenmerk dat slechts een klein effect sorteert, ook al zijn ze beide significant.

Als er meer studies naar hetzelfde boodschapkenmerk zijn uitgevoerd, wordt het mogelijk om middels een meta-analyse een schatting te maken van de effectgrootte. In een meta-analyse worden de resultaten van alle studies die zijn uitgevoerd om eenzelfde vraag te beantwoorden op statistisch verantwoorde wijze gecompileerd. In grote lijnen komt het erop neer dat men (1) alle (gepubliceerde én ongepubliceerde) studies probeert te vinden waarin de overtuigingskracht van eenzelfde communicatiestrategie is onderzocht, (2) voor elk van die studies een effectgrootte berekent (rekening houdend met bijvoorbeeld de betrouwbaarheid van de schaal waarmee een variabele is gemeten) en (3) een gemiddelde effectgrootte berekent (rekening houdend met de omvang van de steekproef) (zie voor uitgebreidere besprekingen Eisend, 2017; Lipsey & Wilson, 2001).

In de volgende paragrafen zal ik de vragen die ik in de inleiding heb gesteld, beantwoorden op basis van meta-analyses. Ik rapporteer daarbij zo veel mogelijk de volgende gegevens:

- 1 de meta-analytische effectgrootte (r): ik rapporteer de correlatiecoëfficiënt. Deze varieert van -1 tot 1 waarbij een min-teken uitdrukt dat het om een negatief effect gaat. Hoe dichterbij 1 (of -1) ligt, hoe sterker het effect. Cohen (1992) karakteriseert een r van .1 als een klein effect, .3 als een medium effect en .5 als een groot effect.
- 2 het aantal effectgroottes (K): de K geeft het aantal effectgroottes waarop de meta-analytische effectgrootte is gebaseerd.
- 3 het aantal proefpersonen (N): de N geeft aan hoeveel proefpersonen betrokken waren bij de schatting van de effectgrootte. Dit aantal ontbreekt in sommige meta-analyses, met name als het een deelverzameling van de geanalyseerde studies betreft.

O'Keefe (2017) laat zien dat het interpreteren van effectgroottes regelmatig misgaat. In veel meta-analyses van onderzoek naar het overtuigingsproces wordt de overtuigingskracht van de ene communicatiestrategie vergeleken met die van een andere. De effectgrootte geeft daarmee antwoord op de vraag: in welke mate is deze communicatiestrategie *overtuigender* dan die communicatiestrategie? In veel gevallen wordt de effectgrootte echter geïnterpreteerd als 'deze strategie is wel overtuigend en die niet'.

Een van de voorbeelden die hij aanhaalt, is de meta-analyse van Tannenbaum et al. (2015) van onderzoek naar angstaanjagende boodschappen. In de methodeparagraaf staat: "Overall, all results should be interpreted as the effect of exposure to messages designed to depict relatively high levels of fear compared to conditions designed to depict relatively lower levels of fear (including no fear)" (Tannenbaum et al., p. 1183). In de conclusie stellen ze echter "Fear appeals are effective" (p. 1196). O'Keefe (2017) wijst erop dat de conclusie moet luiden: sterk angstaanjagende boodschappen zijn overtuigender dan zwakkere angstaanjagende boodschappen. In absolute zin zouden beide versies van angstaanjagende boodschappen overtuigend kunnen zijn (maar sterk angstaanjagende boodschappen in sterkere mate). Bij de interpretatie van de effectgroottes is het dus van belang om te weten wat de controleconditie was op basis waarvan de effectgrootte is berekend. Ik geef daarom bij elke meta-analyse aan wat die controle-conditie was.

Hoe beïnvloedbaar is gedrag?

Met communicatieve interventies wordt geprobeerd het gedrag van mensen te beïnvloeden door bepaalde gedachten of gevoelens op te roepen. Zo gebruikt de overheid argumenten om mensen ertoe aan te zetten hun huis

te isoleren, wordt in gezondheidscommunicatie mensen angst aangejaagd om ze meer te laten bewegen, spelen goede-doelenorganisaties in op ons schuldgevoel en proberen reclamemakers ons te laten lachen. Als gevolg van de zo beïnvloede gedachten en/of gevoelens zouden mensen hun intentie aanpassen, wat vervolgens tot gedragsverandering moet leiden. In onderzoek wordt de effectiviteit van een interventie vaak afgeleid uit het effect op de intentie of op de attitude.

In eerste instantie lijkt de focus op de intentie gerechtvaardigd. Sheeran (2002) geeft een overzicht van verschillende meta-analyses van onderzoek naar de relatie tussen intentie en gedrag en rapporteert een sterk verband: $r = .53$ ($K = 422$, $N = 82.107$). Manning (2009) rapporteert een vergelijkbare relatie: $r = .56$, ($K = 165$, $N = 46.676$). Aan de andere kant laten deze effecten ook zien dat er een behoorlijke kloof bestaat tussen de intentie en het gedrag. Sheeran (p. 7) laat zien dat 47% van de mensen die van plan waren om een bepaald gedrag uit te voeren dat uiteindelijk niet doet.

Webb en Sheeran (2006) betogen dat deze resultaten mogelijk een overschatting vormen van de relatie tussen intentie en gedrag. Als mensen gevraagd wordt naar hun intentie, dan kunnen ze hun daadwerkelijk gedrag als indicatie voor die intentie gebruiken: 'Ik heb dit gedrag uitgevoerd. Dan zal ik dat ook wel van plan zijn geweest.' In dat geval is de intentie geen voorspeller van gedrag, maar het gedrag een voorspeller van de intentie. Om een scherper beeld te krijgen van de relatie tussen intentie en gedrag, voeren Webb en Sheeran een meta-analyse uit van studies waarbij de intentie succesvol is gemanipuleerd door een interventie en vervolgens het daadwerkelijk gedrag is geobserveerd. De redenering is dat als de intentie inderdaad een voorspeller is van het gedrag, een interventie waarmee het lukt om die intentie te veranderen óók een (vergelijkbare) verandering in gedrag zou moeten opleveren.

In hun meta-analyse vinden Webb en Sheeran (2006) een gemiddeld interventie-effect van $r = .31$ ($K = 47$, $N = 8.802$) op intentie. Het bijbehorende effect op gedrag blijkt veel kleiner: $r = .18$ ($K = 47$, $N = 8.802$). Een intentieverandering blijkt dus geen garantie voor een vergelijkbare gedragsverandering. De mate waarin dat het geval is, blijkt mede afhankelijk van het type gedrag. Webb en Sheeran onderscheiden gewoontegedrag van andere vormen van gedrag. Gewoontegedrag betreft gedragingen die herhaaldelijk in dezelfde omstandigheden worden uitgevoerd en na verloop van tijd gedachteloos worden uitgevoerd (zoals het omdoen van een autogordel) (Wood & Rünger, 2016). De invloed van intentieveranderingen op gewoontegedrag is kleiner ($r = .11$, $K = 12$, $N = 1.992$) dan op andere gedragsvormen ($r = .35$, $K = 35$, $N = 6.810$).

De gebrekkige voorspelling van gewoontegedrag door de intentie komt overeen met de heersende ideeën over bewust en onbewust gedrag: “Behavior is normally carried out by unconscious, automatic processes, while consciousness can occasionally intervene to override, regulate, redirect, and otherwise alter the stream of behavior – often at a distance, with unconscious processes filling in” (Baumeister & Bargh, 2014, p. 37). De auteurs vergelijken het onbewuste met een automobilist en het bewustzijn met een navigatiesysteem. In veel gevallen weten automobilisten waar zij heen willen (bijvoorbeeld kantoor, supermarkt) en hebben ze voor het bereiken ervan geen navigatiesysteem nodig. Voor het bezoeken van net verhuisde vrienden zal een chauffeur wel het navigatiesysteem raadplegen, maar ook in dat geval stuurt de chauffeur, niet het navigatiesysteem.

Op vergelijkbare wijze stuurt het onbewuste het leeuwendeel van het menselijk gedrag, zeker als het om herhaald gedrag in een vergelijkbare situatie gaat. We bedenken niet elke keer wat we bij het ontbijt eten en of we met de auto of het openbaar vervoer naar het werk gaan. Er wordt pas op gedragsalternatieven gereflecteerd als een situatie nieuw is (je verandert van baan en moet besluiten hoe je op het nieuwe adres komt) of tegenstrijdige doelen activeert (de weegschaal geeft aan dat je zwaarder bent geworden maar de croissant bij het ontbijt ziet er zo lekker uit). Communicatieve interventies waarmee geprobeerd wordt dit gedrag te beïnvloeden, kunnen wellicht intenties veranderen. Omdat gewoontegedrag echter vooral wordt opgeroepen door de situatie, blijft het gedrag vaak onveranderd.

Interessant in dit verband is de vorming van implementatie-intenties om de intentie-gedragskloof te overbruggen. Een implementatie-intentie is een ‘als-dan’-plan waarin wordt aangegeven wanneer, waar en hoe het gedrag uitgevoerd zal worden (Gollwitzer, 1999). Gollwitzer en Sheeran (2006) voerden een meta-analyse uit van onderzoek naar het effect van de vorming van implementatie-intenties op het uiteindelijk gedrag. Ze vergeleken het gedrag van proefpersonen die de intentie hadden gevormd om een bepaald gedrag uit te voeren met dat van proefpersonen die daarnaast hadden gepland hoe en wanneer ze dat gedrag uit wilden voeren. Vorming van implementatie-intenties bleek de kans te verhogen dat het gedrag werd uitgevoerd ($r = .31$, $K = 94$, $N = 8.461$). Juist door een intentie te koppelen aan elementen van de situatie lijkt het onbewuste die intentie te betrekken bij het gedrag.

Welke factoren bepalen de intentie?

Communicatieve interventies richten zich vaak niet rechtstreeks op het veranderen van de intentie, maar op determinanten die daaraan ten grondslag liggen. In verschillende theoretische modellen (bijvoorbeeld Fishbein & Ajzen, 2010; Fishbein & Yzer, 2003) worden de volgende determinanten onderscheiden: de evaluatie van het gedrag (de attitude), de mate waarin het gedrag door anderen wordt uitgevoerd (de descriptieve norm) of als gewenst wordt beschouwd (de injunctieve norm) en de mate waarin men zichzelf in staat acht het gedrag uit te voeren (waargenomen gedragscontrole, eigen-effectiviteit).

Er is inmiddels een groot aantal meta-analyses uitgevoerd van onderzoek naar de relaties tussen deze attitude, norm en waargenomen controle enerzijds en intentie (en gedrag) anderzijds. Sheeran et al. (2016) geven een overzicht van 18 van dergelijke meta-analyses en rapporteren de volgende correlaties voor de relatie met de intentie (attitude: $r = .54$, $K = 501$; norm: $r = .37$, $K = 433$; waargenomen controle: $r = .51$, $K = 455$) en met gedrag (attitude: $r = .33$, $K = 585$; norm: $r = .20$, $K = 391$; waargenomen controle: $r = .28$, $K = 503$).

Voor deze correlaties geldt hetzelfde bezwaar als bij de schatting van de correlatie tussen intentie en gedrag: bij het beantwoorden van de attitudevragen (of die over de norm of waargenomen controle) kunnen proefpersonen hun reactie (mede) baseren op hun intentie of gedrag. Om die reden voeren Sheeran et al. (2016) een meta-analyse uit waarbij ze eerst vaststellen of een interventie erin is geslaagd om één van de intentiedeterminanten te beïnvloeden, om vervolgens te kijken of er een vergelijkbaar effect optreedt bij de intentie en bij het gedrag. Tabel 1 bevat een overzicht van hun resultaten. In de eerste kolom staat de gemiddelde effectgrootte van de interventie op de desbetreffende determinant (dus: in welke mate is het met de interventie gelukt om de attitude, norm, enzovoort te veranderen). In de kolommen ernaast staat wat vervolgens het interventie-effect is op de intentie en op het gedrag.

Tabel 1 Gemiddelde effectgroottes van de interventie op gedragsdeterminanten, intentie en gedrag

Determinant	Interventie-effect	Intentie-effect	Gedragseffect
Attitude	$r = .23$ $K = 87$, $N = 34.993$	$r = .23$ $K = 59$, $N = 15.145$	$r = .19$ $K = 67$, $N = 31.328$
Norm	$r = .30$ $K = 21$, $N = 10.087$	$r = .24$ $K = 16$, $N = 6.039$	$r = .18$ $K = 17$, $N = 9.337$
Eigen-effectiviteit	$r = .31$ $K = 109$, $N = 36.593$	$r = .25$ $K = 50$, $N = 12.450$	$r = .23$ $K = 90$, $N = 29.520$

De resultaten laten zien dat beïnvloeding van de determinanten gevolgen heeft voor intentie en gedrag maar ook dat die effecten kleiner zijn dan de in eerdere meta-analyses gerapporteerde relaties. Sheeran et al. (2016) rapporteren een mediatie-analyse waaruit blijkt dat het effect van de interventie op gedrag deels wordt gemedieerd door de intentie. Samenvattend laat deze analyse zien dat het zinvol is om te proberen de attitude, norm en eigen-effectiviteit te beïnvloeden omdat dit leidt tot de beoogde veranderingen in intentie en gedrag. Tegelijkertijd laat deze analyse zien dat die veranderingen in intentie relatief beperkt zijn en dat die beperking in nog sterkere mate voor gedrag geldt.

In hoeverre beïnvloeden communicatie-interventies gedrag?

Er zijn verschillende manieren om gedragsdeterminanten te beïnvloeden. In de context van gezondheidsgedrag noemen Steinmetz, Knappstein, Ajzen, Schmidt en Kabst (2016) er bijvoorbeeld dertien, waaronder financiële prikkels en steun in de sociale omgeving. De drie meest voorkomende strategieën zijn echter het geven van informatie, het versterken van vaardigheden, persuasieve communicatie (gericht op het overtuigen van de doelgroep dat ze het gedrag uit kunnen voeren) en de vorming van implementatie-intenties; het geven van informatie wordt vaak gecombineerd met andere strategieën. Interessant in de analyse van Steinmetz et al. is de poging om de relatieve impact van de verschillende strategieën met elkaar te vergelijken waarbij ze het effect van de strategie vergelijken met een controlegroep die geen interventie of een placebo-interventie krijgt. In vergelijking daarmee blijkt persuasieve communicatie een klein tot medium effect op de intentie te sorteren ($r = .17$, $K = 10$), maar dat effect blijkt groter dan dat van vorming van implementatie-intenties ($r = .05$, $K = 16$), terwijl het versterken van vaardigheden geen significant effect op de intentie laat zien.

Anker, Feeley, McCracken en Lagoe (2016) voerden een meta-analyse uit naar de invloed van massamediale gezondheidscampagnes op gedrag. Zij namen uitsluitend studies mee waarin zowel in de experimentele conditie als in de controleconditie een pretest én een posttest waren afgenomen en waarbij de proefpersonen in de controleconditie vergelijkbare proefpersonen waren uit een andere regio, die daardoor niet waren blootgesteld aan de campagne. Voor zowel de experimentele als de controlegroep werd vervolgens een verschilscore berekend op basis van de post- en pretestscores.

Zij vonden een overall effect van $r = .054$ ($K = 61$, $N = 164.949$), dat echter varieerde voor verschillende gedragingen. Voor roken, seksueel gedrag en kankergerelateerd gedrag werd geen gedragseffect gevonden, campagnes voor meer bewegen ($r = .052$, $K = 5$) en ter voorkoming van hartaandoeningen ($r = .047$, $k = 11$) sorteerden kleine effecten, terwijl er grotere effecten werden gevonden voor veilig-verkeer- ($r = .15$, $K = 8$) en voedingscampagnes ($r = .189$, $K = 3$).

Feeley en Moon (2009) voerden een meta-analyse uit van onderzoek naar de effectiviteit van orgaandonatiecampagnes waarbij het zich registreren als donor en het op de hoogte brengen van familieleden van dat besluit de belangrijkste gedragsdoelen waren. In deze meta-analyse werden 23 studies met in totaal meer dan 6 miljoen deelnemers meegenomen ($N = 6.647.801$). De scores van de groep die blootgesteld was aan een campagne werden vergeleken met de scores op een voormeting van diezelfde groep, met een controlegroep die geen boodschap kreeg, of met een controlegroep die een boodschap kreeg over een ander thema. De gemiddelde effectgrootte voor registratiegedrag bedroeg ($r = .05$, $K = 19$) en was daarmee vergelijkbaar met die voor het bespreken van de beslissing met familieleden ($r = .05$, $K = 15$).

In de bovengenoemde meta-analyses gaat het om communicatie-interventies in de vorm van betogen met een duidelijke persuasief doel zoals brochures of posters. Er worden echter ook interventies in de vorm van verhalen ingezet om gezondheidsgedrag te beïnvloeden. Shen, Sheer en Li (2015) brengen de impact van narratieve interventies in kaart. Zij vergelijken daarbij het effect van deze interventies met niet-narratieve interventies. Zij rapporteren een effect van $r = .063$ ($K = 34$, $N = 9.330$) waarbij ze effecten op attitudes, intenties en gedrag combineren. Ook zij vinden verschillen in effectgroottes voor verschillende gedragingen. Waar narratieve interventies meer invloed hebben dan non-narratieve interventies op preventiegedrag zoals het gebruik van zonnebrandmiddel ($r = .075$, $K = 20$) en detectiegedrag zoals borstzelfonderzoek ($r = .091$, $K = 9$), sorteren narratieve interventies gericht op het stoppen met een bepaald gedrag (bijvoorbeeld roken) een vergelijkbaar effect als niet-narratieve interventies ($r = -.008$, $K = 5$).

Braddock en Dillard (2016) voerden ook een meta-analyse uit van onderzoek naar de effecten van narratieve interventies, waarbij ze zich niet beperkten tot gezondheidscommunicatie. De controlecondities in hun meta-analyse zijn divers: voormetingen, geen boodschap, een ongerelateerde boodschap of (voor- en/of na-)metingen bij een vergelijkbare of minder vergelijkbare groep. Dit maakt de interpretatie van de effectgrootte ingewikkeld. Braddock en Dillard onderscheiden effecten op overtuigingen

($r = .17$, $K = 37$, $N = 7.376$), attitudes ($r = .19$, $K = 40$, $N = 7.132$), intenties ($r = .17$, $K = 28$, $N = 5.211$) en gedrag ($r = .23$, $K = 5$, $N = 978$). De gemiddelde effectgroottes zijn hoger dan die gerapporteerd door Shen et al. (wat veroorzaakt kan zijn door de verschillen in controlegroepen tussen beide meta-analyses), maar zijn nog steeds 'klein' tot 'medium'.

Eisend en Tarrahi (2016) voerden een zogenaamde meta-meta-analyse uit van onderzoek naar de impact van reclame op de attitudes, intenties en het gedrag van consumenten. Bij een meta-meta-analyse vormen niet de schattingen van effectgroottes in individuele studies de onderzoekseenheden, maar de schattingen van effectgroottes uit eerder uitgevoerde meta-analyses. In hun analyse vinden zij vergelijkbare effecten voor attitudes en intenties. Ze rapporteren daarom een samengestelde maat. Voor deze attitude/intentie-maat rapporteren ze een effect van reclame van $r = .21$, gebaseerd op 119 meta-analytische effectgroottes (K_{m-a}), die op hun beurt gebaseerd zijn op 2.190 effectgroottes waarbij 403.430 proefpersonen waren betrokken. Voor de invloed op consumentengedrag vinden ze een vergelijkbaar effect ($r = .20$, $K_{m-a} = 44$, $K = 2.191$, $N = 688.990$). Een complicerende factor voor de interpretatie is dat de controlegroepen in deze studies meestal een alternatieve versie van de reclameboodschap hebben gezien, bijvoorbeeld een reclame zonder humor in onderzoek naar humor in reclame of een boodschap met een onbekend persoon in een studie naar het effect van beroemdheid. De effectgroottes zeggen dus iets over de vraag wat variaties op dit soort dimensies voor verschillen in effect bewerkstelligen, maar niet over de vraag of reclame effect heeft.

De verschillende meta-analyses laten een vergelijkbaar patroon zien. De effecten van communicatie op gedrag zijn non-existent of klein en blijken mede afhankelijk van het type gedrag en mogelijk van het soort interventie. Op basis van de meta-analyse van Braddock en Dillard (2016) lijken de narratieve interventies een sterker effect te sorteren dan de retorische interventies. Daarbij moet worden opgemerkt dat het aantal studies (5) en proefpersonen (978) relatief klein is en het onduidelijk is wat in deze gevallen de controleconditie was.

Overtuigingsstrategieën: de impact van ethos, pathos en logos

In de klassieke retorica worden drie overtuigingsmiddelen onderscheiden: het karakter van de bron (ethos), de gevoelens van het publiek (pathos) of

de kracht van argumenten (logos). Voor de effectiviteit van elk van deze overtuigingsmiddelen zijn verschillende meta-analyses uitgevoerd.

De geloofwaardigheid van de bron

Er zijn veel studies gedaan naar de invloed van bronkenmerken op overtuigingskracht. Kumkale, Abarracín en Seignourel (2010) voerden een meta-analyse uit waarbij ze experimenten selecteerden waarin (1) de geloofwaardigheid van de bron succesvol was gemanipuleerd en (2) de overtuigingskracht niet alleen direct ná blootstelling aan de boodschap maar ook (gemiddeld twee weken) later werd gemeten. Op die manier biedt hun analyse zicht op de stabiliteit van een eventueel effect. Daarnaast onderscheidden Kumkale et al. proefpersonen die al een attitude hadden over dit onderwerp van proefpersonen die dat nog niet hadden én gingen ze na of de proefpersonen wel of niet in de gelegenheid waren om zich een oordeel te vormen over het thema op basis van de inhoud van de boodschap.

Onmiddellijk na blootstelling aan de boodschap leidde een geloofwaardiger bron tot een sterkere acceptatie van het standpunt in vergelijking met een minder geloofwaardige bron ($r = .21$, $K = 53$, $N = 4.502$). Bij het latere meetpunt was dit effect bijna gehalveerd, maar nog steeds significant ($r = .11$, $K = 53$, $N = 4.502$). Het sterkste effect van de geloofwaardigheid van de bron werd gevonden als de proefpersonen nog geen standpunt hadden ingenomen én over weinig kennis over het onderwerp beschikten ($r = .35$, $K = 16$). Ook hier is het effect bij de latere meting gehalveerd ($r = .17$, $K = 16$).

In de al eerder genoemde meta-meta-analyse van Eisend en Tarrahi (2016), wordt een schatting gemaakt van de effectgrootte voor bronkenmerken in reclame. Daarbij gaat het niet alleen om manipulatie van de geloofwaardigheid maar ook om bronkenmerken zoals fysieke aantrekkelijkheid of beroemdheid. Zij rapporteren een effect op de attitude/intentie ($r = .287$, $K_{m-a} = 11$, $K = 290$, $N = 23.005$). Daarnaast rapporteren zij een effect van bronkenmerken op consumentengedrag ($r = .217$, $K_{m-a} = 18$, $K = 1.150$, $N = 304.661$). Uit beide meta-analyses komt dus naar voren dat bronkenmerken invloed kunnen hebben op de overtuigingskracht.

De gevoelens van het publiek

Angie, Connelly, Waples en Kligyte (2011) voerden een meta-analyse uit naar de invloed van emoties op oordeelsvorming (bijvoorbeeld over de waarschijnlijkheid van een bepaalde gebeurtenis) en op voorkeuren (bijvoorbeeld voor een arts of een bepaalde baan). Zij voeren aparte meta-analyses uit voor onderzoeken waarin het effect van een emotie werd vergeleken met een controlegroep en voor onderzoeken waarin dat effect werd

vergeleken met het effect van een andere emotie (bijvoorbeeld boosheid versus angst).

Angie et al. (2011) waren *niet* geïnteresseerd in de vraag of boosheid altijd een positief (of negatief) effect had, maar ze wilden eenvoudigweg weten *of* het een effect had. Als in de ene studie boosheid tot een positiever oordeel leidde (bijvoorbeeld $r = .20$) en in de andere studie tot een negatiever oordeel ($r = -.15$), dan werden de absolute effectgroottes bij elkaar opgeteld ($r = .20 + r = .15$), omdat ze geïnteresseerd waren in de vraag of emoties invloed hebben en niet zozeer *wát* de richting is. Het aantal effectgroottes en proefpersonen is relatief beperkt. In Tabel 2 staan de gemiddelde (absolute) effectgroottes voor de verschillende emoties uitgesplitst naar oordeelsvorming en voorkeuren voor de onderzoeken waarin het effect van de emotie is vergeleken met een controlegroep.

Tabel 2 Effectgroottes van specifieke emoties voor oordeelsvorming en voorkeuren in Angie et al. (2011)

Emotie	Oordeelsvorming			Voorkeuren		
	r	K	N	r	K	N
Boosheid	.03 (n.s.)	6	535	.13	10	862
Angst	.05	6	542	.09	6	484
Verdriet	.09	9	553	.16	13	975
Geluk	.08	5	304			
Afkeer				.18	2	192
Schuldgevoel				.44	3	333

Er zijn ook andere meta-analyses uitgevoerd van onderzoek naar specifieke emoties. Met name ‘angst’ is een vaak onderzocht fenomeen. De meest recente – en omvangrijke – meta-analyse is uitgevoerd door Tannenbaum et al. (2015). Zij rapporteren een positief effect van angstaanjagende boodschappen in vergelijking met minder angstaanjagende boodschappen op een gecombineerde maat van attitude, intentie en gedragsmaten ($r = .14$, $K = 248$, $N = 27.372$). Daarnaast vinden ze grotere effecten voor boodschappen die meer angst aanjagen als die boodschappen ook informatie bevatten over de uitvoerbaarheid en effectiviteit van het aanbevolen gedrag ($r = .21$, $K = 92$), als die boodschap ook informatie bevat ter onderbouwing van de ernst en waarschijnlijkheid van het gevaar ($r = .19$, $K = 78$) en als het gaat om eenmalig gedrag (bijvoorbeeld vaccinatie) ($r = .21$, $K = 82$).

Opvallend in het overzicht van Tabel 2 is de hoge score voor schuldgevoel. Xu en Guo (2018) voerden een meta-analyse uit voor het vaststellen van het verband tussen schuldgevoel en een gecombineerde attitude/intentie-maat. De gerapporteerde (forse) effectgrootte, $r = .49$ ($K = 8$,

$N = 2.061$) geeft dus de gewogen gemiddelde correlatiecoëfficiënt weer. Het is geen schatting van het effect van een boodschap die schuldgevoelens oproept in vergelijking met een boodschap die dat niet, of in mindere mate doet. Hier kan dus sprake zijn van een overschatting omdat mensen hun (positieve) intentie mogelijk als indicatie gebruiken voor de mate waarin ze een schuldgevoel hebben ervaren.

Ook in reclamecontexten blijkt het oproepen van emoties de overtuigingskracht te verhogen. Pham, Geuens en De Pelsmacker (2013) beschikten over de attitudes ten opzichte van de commercial en de attitudes ten opzichte van het aangeprezen merk voor een corpus van 1070 televisiecommercials gegenereerd door een grote (1576) en heterogene steekproef van consumenten die gemiddeld 20 tot 50 commercials hadden beoordeeld. Elke commercial werd daarnaast beoordeeld door een groep van 12 tot 14 (onafhankelijke) beoordelaars op de mate waarin de commercial warme gevoelens, opwinding en blijdschap teweegbracht. De mate waarin een commercial emotionele reacties oproep, bleek significant samen te hangen met zowel de attitude ten opzichte van de advertentie ($r = .50$) als de attitude ten opzichte van het merk ($r = .26$). Op basis van regressieanalyses laten Pham et al. zien dat de door de commercial opgeroepen gevoelens zowel direct als indirect (via de attitude ten opzichte van de commercial) de attitude ten opzichte van het merk beïnvloedden, waarbij de indirecte invloed groter was dan de directe. De metingen van de attitude ten opzichte van de commercial en die ten opzichte van het merk zijn door dezelfde proefpersonen gegenereerd, wat voor een sterkere relatie tussen de twee attitudes kan hebben gezorgd.

Een veel voorkomende reclamestrategie is het gebruik van humor om vrolijkheid op te roepen. Eisend (2011) voerde een meta-meta-analyse uit naar de invloed van humor in reclame. Hij rapporteert een effect van humor op de attitude ten opzichte van het merk ($r = .095$, $K = 49$, $N = 4.298$) en een (sterker) effect op de attitude ten opzichte van de advertentie ($r = .192$ ($K = 87$, $N = 18.314$)). Eisend voert vervolgens een pad-analyse uit waarbij hij de attitude ten opzichte van de advertentie als belangrijkste voorspeller van de attitude ten opzichte van het merk identificeert. In combinatie met de resultaten van Pham et al. (2013) suggereert dit dat het oproepen van emoties tot een positievere attitude ten opzichte van de boodschap leidt, wat weer leidt tot een positievere attitude ten opzichte van het merk.

Wirtz, Sparks en Zimbres (2018) voerden een meta-analyse uit naar de effectiviteit van het appelleren aan seksuele gevoelens in reclame. Het gebruik van deze strategie had een negatief effect op de attitude ten opzichte van het merk ($r = -.11$, $K = 47$, $N = 4.152$) maar geen invloed op de intentie om

het product aan te schaffen ($r = -.01$, $K = 54$, $N = 6.964$). Ook de attitude ten opzichte van de advertentie leek niet gevoelig voor het gebruik van deze strategie ($r = -.04$, $K = 100$, $N = 11.497$) maar bij het opsplitsen naar gender bleek deze bij vrouwen tot een lagere ($r = -.19$, $K = 35$, $N = 3.319$) en bij mannen tot een hogere waardering voor de advertentie ($r = .13$, $K = 28$, $N = 2.037$) te leiden.

De kwaliteit van argumenten

De laatste overtuigingsstrategie is het gebruik van argumenten. In het Elaboration Likelihood Model (ELM; Petty & Cacioppo, 1986) wordt voorspeld dat als mensen in staat en gemotiveerd zijn om een correct oordeel te vormen over een standpunt, de kwaliteit van de aangevoerde argumenten doorslaggevend is voor de acceptatie van het standpunt. Is men niet in staat om de argumenten zorgvuldig af te wegen of ontbreekt de motivatie daartoe, dan laat men zich leiden door perifere cues.

Carpenter (2015) heeft een meta-analyse uitgevoerd van onderzoeken waarin het effect van argumentkwaliteit op de attitude is onderzocht. Het gaat daarbij om 134 effectgroottes voor de mate waarin een boodschap met sterke argumenten tot een positievere attitude leidt dan een boodschap met zwakke argumenten en waarbij 15.426 proefpersonen zijn betrokken. Hij rapporteert geen overall effect van argumentkwaliteit, maar splitst het effect uit voor proefpersonen die de argumenten kritischer zouden bekijken en proefpersonen die ze meer oppervlakkig zouden verwerken. Voor de kritische groep rapporteert hij een sterk effect van argumentkwaliteit ($r = .47$, $K = 134$). Ook voor de meer oppervlakkig verwerkende proefpersonen vindt hij dat een boodschap met sterke argumenten tot een positievere attitude leidt dan een boodschap met zwakke argumenten, al is dit effect kleiner ($r = .14$, $K = 134$). Interessant is ook een tweede vergelijking die Carpenter rapporteert. Bij sterke argumenten is de attitude van de kritisch verwerkende proefpersonen positiever dan die van oppervlakkig verwerkende proefpersonen ($r = .19$, $K = 53$); bij zwakke argumenten is die attitude juist negatiever ($r = -.20$, $K = 39$).

In eerste instantie lijken deze resultaten te bevestigen dat argumenteren een effectieve overtuigingsstrategie is. Voor kritisch verwerkende proefpersonen is het effect zelfs groot. Toch roepen de resultaten ook vragen op. Waarom zouden oppervlakkig verwerkende proefpersonen óók gevoelig zijn voor argumentkwaliteit? Veel onderzoek naar redeneerprocessen laat juist zien dat mensen slechts onder zeer bepaalde omstandigheden hun kritische vermogens aanspreken (Kahneman, 2011). Daarnaast is het opvallend dat zwakke argumenten tot negatievere attitudes leiden. Waarom zou

een zwak argument vóór een standpunt tot sterkere verwerping van dat standpunt leiden?

Een mogelijke verklaring voor deze resultaten vormt de wijze waarop argumentkwaliteit wordt gemanipuleerd. Raju, Unnava en Montgomery (2009) maakten in hun onderzoek gebruik van het volgende argument voor een auto: 3 jaar volledige garantie, scoort 4 op een 5 puntschaal voor veiligheid en rijdt 9,5 kilometer op 1 liter benzine in de stad. Is dit nu een sterk of een zwak argument? Volgens Raju et al. is dit een zwak argument en dat is ook logisch als je het vergelijkt met het sterke argument dat zij gebruikten: 5 jaar volledige garantie, scoort 5 op een 5 puntschaal voor veiligheid en rijdt 13,5 kilometer op 1 liter benzine in de stad. Proefpersonen zien echter maar één van de twee argumenten en dan lijkt het toch lastig om het zwakke argument eenduidig als zwak te zien.

Er zijn echter studies waarbij het gemakkelijk is om een ‘zwak’ argument te identificeren als zwak. Calanchini, Moons en Mackie (2016) gebruikten bijvoorbeeld als argument voor belastingverhoging dat de extra inkomsten zouden worden gebruikt om snelwegen te verfraaien zodat lifters beter opvielen en gemakkelijker beboet konden worden. Tormala, Briñol en Petty (2006) prezen een wasmiddel aan vanwege het prachtige design van de verpakking, wat het mogelijk maakte om het als kunstobject te gebruiken. Chu en Kamal (2008) gebruikten een advertentie waarin een laptop werd ‘aangeprezen’ vanwege zijn verouderde technologie, middelmatige kwaliteit van de onderdelen, zwaarste materialen, zwakste beeldkwaliteit en gemiddelde accuduur.

Zoals uit deze voorbeelden blijkt, is ‘manipulatie van argumentkwaliteit’ een parapluterm voor uiteenlopende manipulaties. Vanuit een normatief perspectief kan het daarbij gaan om daadwerkelijk zwakke argumenten (zoals in het voorbeeld van de advertentie voor een auto), maar ook om irrelevante informatie (zoals in het voorbeeld waarin de snelwegverfraaiing bepleit wordt) of zelfs om (sterke) tegenargumenten (zoals in de advertentie voor een laptop). In die laatste twee gevallen is het voorstelbaar dat ook oppervlakkig verwerkende proefpersonen zien dat het hier om argumenten gaat die geen ondersteuning vormen voor het standpunt en er dus ook bij hen een effect van argumentkwaliteit optreedt. Het gebruik van tegenargumenten zou kunnen verklaren waarom boodschappen met dit soort argumenten tot negatievere attitudes leiden. Weerman, Hoeken en Hornikx (2016) laten voor een kleine set studies zien dat dit inderdaad het geval kan zijn. In elk geval zijn er aanwijzingen dat de door Carpenter (2015) gevonden effectgrootte voor argumentkwaliteit een overschatting vormt van het daadwerkelijke effect.

Effecten van boodschaptactieken

Er zijn ook verschillende meta-analyses uitgevoerd naar boodschaptactieken. Grofweg kunnen die worden ingedeeld in argumentatie-gerelateerde tactieken (evidentietype, één- of tweezijdige boodschappen en al dan niet expliciet trekken van conclusies) en meer vorm-gerelateerde keuzes zoals *gain* versus *loss framing*, metaforen en het gebruik van afbeeldingen. In Tabel 3 staan de resultaten voor deze meta-analyses.

Tabel 3 Resultaten voor meta-analyse voor verschillende boodschaptactieken

Bron	Boodschapkenmerk	Effect
Zebregs et al. (2014)	Narratieve vs. Statistische evidentie op intentie	$r = .05$ ($K = 7, N = 1.338$)
O'Keefe (2002)	Conclusie expliciet of impliciet	$r = .10$ ($K = 35, N = 14.215$)
Eisend (2006)	Ook nadelen of alleen voordelen op merkattitude:	$r = .11$ ($K = 69, N = 3.525$)
	op intentie:	$r = .12$ ($K = 41, N = 3.019$)
O'Keefe & Jensen (2007)	Gain versus loss frame bij preventiegedrag	$r = .03$ ($K = 93, N = 21.656$)
O'Keefe & Jensen (2009)	Gain versus loss frame bij detectiegedrag	$r = -.04$ ($K = 53, N = 9.145$)
Sopory & Dillard (2002)	Metaforen	$r = .07$ ($K = 38, N = 3.945$)
Noar et al. (2016)	Afschrikwekkende afbeeldingen op sigarettenpakjes op intentie om te stoppen met roken	$r = .26$ ($K = 8, N = 16.671$)

Noar et al. (2016) vinden een medium effect van afbeeldingen op sigarettenpakjes op de intentie om te stoppen met roken. De overige effectgroottes zijn klein ($.03 \leq r \leq .12$). In absolute zin geeft dit aan hoe beperkt de invloed van boodschapkenmerken is. In relatieve zin zijn de effectgroottes opmerkelijk sterk. In meta-analyses waarin het effect van bijvoorbeeld communicatie-interventies wordt vergeleken met een controlegroep die geen interventie krijgt, zijn de effectgroottes nauwelijks groter, terwijl er bij de boodschaptactieken een vergelijking wordt gemaakt tussen twee dezelfde boodschappen die enkel verschillen op de aanwezigheid van bijvoorbeeld een expliciete conclusie. Ook is het interessant om de effectgrootte van boodschapkenmerken te vergelijken met die van de zogenaamde *compliance gaining* technieken. Feeley, Anker en Aloe (2012) rapporteren een effectgrootte van $r = .13$ ($K = 78, N = 4.892$) voor de *door-in-the-face* techniek, en Dillard, Hunter en Burgoon (1984) melden een effectgrootte van $r = .11$

($K = 42$, $N = 4.927$) voor de *foot-in-the-door* techniek. Deze op het sociale beïnvloedingsmodel van Cialdini (2001) gebaseerde technieken blijken dus ook maar kleine effecten te bewerkstelligen.

Conclusies

Met persuasieve communicatie wordt geprobeerd gedrag te beïnvloeden door bepaalde gedachten of gevoelens bij mensen op te roepen. Deze gedachten en gevoelens moeten leiden tot de vorming van attitudes, normen, en waargenomen gedragscontrole, die op hun beurt leiden tot intenties om het gedrag uit te voeren. Die intentie zou de belangrijkste voorspeller van het gedrag vormen. Onderzoekresultaten laten zien dat deze causale keten veel zwakke schakels kent. Gemiddeld genomen hebben (communicatie- en andere) interventies een klein tot medium effect op de attitude, norm of waargenomen gedragscontrole ($.23 \leq r \leq .31$), resulterend in een iets kleiner effect op de intentie ($.23 \leq r \leq .25$). Een medium effect op de intentie ($r = .31$) leidt over het algemeen tot een klein tot medium effect op gedrag ($r = .18$).

De mate waarin de intentie een voorspeller vormt van het gedrag hangt af van de mate waarin er bewust over het gedrag wordt nagedacht. Veranderingen in intenties voor gedrag dat herhaaldelijk in dezelfde context wordt uitgevoerd, hebben veel minder invloed op het uiteindelijk gedrag ($r = .11$) dan intentieveranderingen voor nieuw of eenmalig gedrag ($r = .35$). De vorming van implementatie-intenties waarin de intentie wordt gekoppeld aan de context waarin dat gedrag moet worden uitgevoerd, vergroot de kans dat de intentie wordt omgezet in gedrag ($r = .31$).

De invloed van communicatieve interventies is klein en varieert voor verschillende soorten gedragingen. Bepaalde vormen van (herhaald) gedrag (bijvoorbeeld roken, alcoholgebruik) lijken ongevoelig voor communicatieve interventies; voor andere gedragingen (meer bewegen, donorregistratie) worden zeer kleine effecten ($r = .05$) gevonden. Iets grotere effecten worden gevonden voor verkeers- en consumentengedrag ($.15 < r < .20$), maar ook die staan nog ver af van wat Cohen (1992) een medium effect zou noemen. Ten slotte zijn er aanwijzingen dat de invloed van narratieve interventies mogelijk iets groter is dan die van retorische interventies.

De drie reeds in de klassieke retorica onderscheiden overtuigingsstrategieën blijken alle in staat om de attitude te beïnvloeden. Schattingen voor het effect van de bron variëren tussen .21 en .29. Ook het oproepen van emoties heeft een beperkte invloed op de overtuigingskracht (humor

in reclame en angst in voorlichting ($.10 < r < .21$). Voor argumentkwaliteit wordt een groot effect ($r = .47$) gevonden op voorwaarde dat de proefpersonen gemotiveerd en in staat zijn om de argumenten te evalueren. Dit is waarschijnlijk een geflatteerd effect omdat er bij de zogenaamde ‘zwakke argumenten’ soms geen sprake is van argumentatie of zelfs van tegenargumenten. De meta-analyses voor de boodschaptactieken laten over het algemeen nog kleinere effectgroottes zien ($.03 \leq r \leq .12$). Een uitzondering is het effect van afbeeldingen op sigarettenpakjes, dat in de buurt van een medium effect komt ($r = .26$).

Deze resultaten leveren een ontvullend beeld op: de effecten van de meeste overtuigingsstrategieën en boodschaptactieken op attitude/intentie zijn (zeer) beperkt. Attitude- en intentieveranderingen blijken op hun beurt maar in beperkte mate te leiden tot gedragsveranderingen. De causale keten van boodschapkenmerk naar gedragsverandering bestaat dus uit zwakke schakels; voorspellingen over gedragseffecten op basis van boodschapkenmerken lijken nauwelijks gerechtvaardigd.

Discussie

In deze paragraaf komen drie vragen aan de orde. Wat zijn de zwakke punten van meta-analyses en welke gevolgen hebben die voor de hier getrokken conclusies? Welke mogelijke consequenties hebben de hier gerapporteerde bevindingen voor de opzet van het onderzoek naar persuasieve communicatie? En ten slotte: welke conclusies kunnen uit deze resultaten worden getrokken voor de adviespraktijk?

Pijnpunten van meta-analyses

De meta-analyse vormt een belangwekkende methode voor het compileren van verschillende studies naar hetzelfde fenomeen, maar er kleven ook nadelen aan. Ten eerste hangt de waarde van de meta-analyse sterk af van de mate waarin de in de analyse opgenomen studies een representatieve afspiegeling vormen van de naar het fenomeen uitgevoerde studies. Het ‘bureaula probleem’ – het gegeven dat elke onderzoeker wel een paar studies heeft uitgevoerd waarover nooit gerapporteerd zal worden – speelt hierbij een belangrijke rol. Bij hun systematische literatuurzoektocht, proberen onderzoekers om naast alle in wetenschappelijke tijdschriften of bundels verschenen onderzoeken ook niet-gepubliceerde onderzoeken te vinden. Daarvoor kijken ze naar dissertaties en conferentieprogramma’s, vragen aan onderzoekers van wél gepubliceerde studies of ze ook nog ongepubliceerde

studies hebben naar hetzelfde fenomeen en plaatsen ze oproepen op relevante mailinglijsten. Toch zijn de meeste in meta-analyses opgenomen studies gepubliceerd in wetenschappelijke tijdschriften of bundels.

In de psychologie is de kans dat een studie gepubliceerd wordt groter als er een effect wordt gerapporteerd dan als er geen effect is gevonden (Ferguson & Heene, 2012). De kans is reëel dat dat ook geldt voor communicatiewetenschappelijk onderzoek. Daardoor bestaat het gevaar van een oververtegenwoordiging van studies waarin een effect is gevonden. In de meta-analyses wordt als reactie hierop vaak een *Failsafe N* gerapporteerd. Dit is een schatting van het aantal studies met een niet-significant effect dat nodig zou zijn om het in de meta-analyse gevonden overall effect te reduceren van significant tot niet-significant. Dat aantal is meestal behoorlijk indrukwekkend. Tegelijkertijd reduceert de *Failsafe N* daarmee de uitkomst van de meta-analyse tot een antwoord op de vraag: Is het effect significant of niet? Daarmee vervalt het voordeel van een meta-analyse dat die een indicatie kan geven van de effectgrootte. Als studies waarbij een significant effect wordt gerapporteerd, oververtegenwoordigd zijn in meta-analyses, dan volgt daar misschien niet uit dat het effect illusoir is, maar wel dat de gevonden effectgrootte een overschatting vormt van het in de populatie bestaande effect. Daar staat tegenover dat in bijna alle onderzoeken het gaat om het effect van een eenmalige aanbieding van een boodschap terwijl zeker bij reclame het uiteindelijke effect ook afhangt van herhaaldelijke blootstelling aan de boodschap.

In de tweede plaats is het de vraag of de in de individuele studies uitgevoerde manipulaties echt betrekking hebben op hetzelfde fenomeen. Meta-analyses zijn gericht op het vaststellen van een effect – een causale relatie tussen een onafhankelijke en een afhankelijke variabele. De wijze waarop de onafhankelijke variabele is gemanipuleerd, kan per studie verschillen en de vraag is in hoeverre elk van die manipulaties een valide operationalisatie vormt van de onafhankelijke variabele. De discussie over de zeer verschillende manieren waarop argumentkwaliteit is gemanipuleerd, illustreert dit probleem. De grote variatie in de aard van de manipulaties geldt niet alleen voor argumentkwaliteit maar bijvoorbeeld ook voor het onderzoek naar narratieve interventies (De Graaf, Sanders, & Hoeken, 2016).

Implicaties voor onderzoek naar persuasieve communicatie

Het overall effect van boodschapkenmerken op attitudes en intenties is eerder klein dan medium of groot. Dit heeft implicaties voor het aantal proefpersonen dat nodig is voor een onderzoek. Bij statistische toetsing geeft de *p*-waarde aan wat de kans is dat we ten onrechte concluderen dat er

sprake is van een effect. Ligt die p -waarde op of boven de kritische grens van .05, dan concluderen we dat er geen effect is. Ook die conclusie kan onterecht zijn: wellicht was er wél een effect. Veel hangt daarbij af van de *power* van de statistische toets: wat is – gegeven de grootte van het effect in de werkelijkheid en het aantal proefpersonen in de verschillende condities – de kans om dat effect in het onderzoek daadwerkelijk te vinden? Daarbij wordt gestreefd naar een power van 80%: als er in het echt een effect is, dan zou dat in vier van de vijf gevallen een statistisch significant verschil moeten opleveren.

Naarmate het werkelijke effect kleiner is, neemt het benodigde aantal proefpersonen om de gewenste power te realiseren toe. Uitgaande van een experiment met twee boodschapversies, een kleine effectgrootte, toetsing door middel van een t -toets voor onafhankelijke waarnemingen tegen een alfa van .05, en een gewenste power van .80, zou elke tekstversie door 393 proefpersonen gelezen moeten worden. Bij eenzijdige in plaats van tweezijdige toetsing, zijn er nog steeds 310 proefpersonen per versie nodig voor een power van .80. In veel onderzoeken ligt het aantal proefpersonen per conditie veel lager, eerder rond de 30 of 40. De power is daardoor ook veel lager: respectievelijk .12 of .14 bij tweezijdige toetsing, .19 of .22 bij eenzijdige toetsing. Zelfs bij het hanteren van een norm van 50 proefpersonen per conditie (zoals nu in sociaal-psychologische experimenten gebruikelijk is), ligt de power nog ver onder het gewenste niveau voor het vinden van een klein effect (namelijk .17 en .26).

Experimenten met onvoldoende proefpersonen vormen een probleem. Stel dat onderzoekers willen weten of een grafiek in kleur meer invloed heeft op de overtuigingskracht dan een in zwartwit, terwijl er in werkelijkheid geen verschil is. Als daar twintig onderzoeksteams onderzoek naar doen met een te kleine steekproef, treedt er gemiddeld één significant effect op. De kans is reëel dat er over dit toevalseffect wel wordt gerapporteerd maar over de andere niet. Preregistratie van studies vormt in dit verband een interessante ontwikkeling omdat daardoor een veel beter zicht ontstaat op het aantal studies dat naar een bepaald fenomeen is uitgevoerd. Een nog interessantere optie zou zijn om weliswaar met steekproeven van 40 proefpersonen per boodschapversie te werken, maar dan wel met tien verschillende teksten die op vergelijkbare manier zijn gemanipuleerd. Ook dat levert de gewenste power op maar biedt tegelijkertijd de mogelijkheid om te controleren of het effect van de manipulatie in al die verschillende teksten vergelijkbaar is.

Implicaties voor advisering over persuasieve communicatie

Adviseurs op het gebied van persuasieve communicatie past bescheidenheid waar het de invloed op gedrag betreft. Met persuasieve communicatie wordt geprobeerd het bewuste ‘navigatiesysteem’ van de mens te (her)programmeren maar dat systeem gebruikt de mens vaak niet. Zelfs als het lukt om met communicatie de attitude of intentie te veranderen, hoeft dat geen gedragsverandering tot gevolg te hebben, omdat dat gedrag meer wordt gestuurd door kenmerken van de situatie dan door de intentie.

Communicatieadviseurs moeten zich bewust zijn van welke gedragingen meer gestuurd worden door intenties (zoals nieuwe of eenmalige gedragingen) en gedragingen waarvoor dat niet, of in veel mindere mate geldt (gewoontegedrag). De kansen om gewoontegedrag te veranderen zijn veel kleiner en veel hangt daarbij af van timing. Ook bij gewoontegedrag zijn er namelijk momenten waarop mensen hun bewuste ‘navigatiesysteem’ raadplegen. Letterlijk is daarvan sprake als een forens altijd met de auto naar het werk gaat maar weer over het openbaar vervoer gaat nadenken als er langdurige wegwerkzaamheden zijn of een nieuw treinstation wordt geopend. Figuurlijk wordt het bewuste ‘navigatiesysteem’ geraadpleegd als iemand elke avond een zak chips wegwerkt op de bank maar zich op 1 januari voorneemt om gezonder te gaan leven. In deze gevallen is het belangrijk om eventuele intentieveranderingen te koppelen aan concrete plannen voor de uitvoering van die intentie.

De al in de oudheid onderscheiden overtuigingsstrategieën – ethos, pathos, logos – lijken ook in de 21^e eeuw nog steeds effectief. Communicatieadviseurs kunnen op basis van de verzamelde inzichten wél iets zeggen over welke boodschaptactieken de reputatie van de bron beschermen of versterken, angst, schuld of vrolijkheid oproepen en welke argumenten meer of minder overtuigend zouden moeten zijn. Daarnaast hebben ze zicht op wat het expliciteren van conclusies tot gevolg heeft en of metaforen bijdragen aan de overtuigingskracht. Tegelijkertijd moet duidelijk zijn dat toepassing van deze inzichten geen garantie vormt voor een overtuigende boodschap. Het zijn kleine rukjes aan het stuur die de kans iets groter maken dat de bestuurder de bocht neemt.

Noot

1. Ik ben Bregje Holleman, Lettica Hustinx, Carel Jansen, Peter Neijens, Daniel O’Keefe, Henk Pander Maat, Peter Jan Schellens en Madelijn Strick zeer dankbaar voor hun commentaar op eerdere versies van dit artikel. Eventuele fouten zijn echter geheel mijn verantwoordelijkheid.

Bibliografie

- Angie, A. D., Connelly, S., Waples, E. P., & Kligyte, V. (2011). The influence of discrete emotions on judgement and decision-making: A meta-analytic review. *Cognition & Emotion*, 25(8), 1393-1422.
- Anker, A. E., Feeley, T. H., McCracken, B., & Lagoe, C. A. (2016). Measuring the effectiveness of mass-mediated health campaigns through meta-analysis. *Journal of Health Communication*, 21(4), 439-456.
- Baumeister, R. F., & Bargh, J. A. (2014). Conscious and unconscious. In J. W. Sherman, B. Gawronski, & Y. Trope (Eds.), *Dual-process Theories of the Social Mind* (pp. 35-49). New York, NY: Guilford Press.
- Braddock, K., & Dillard, J. P. (2016). Meta-analytic evidence for the persuasive effect of narratives on beliefs, attitudes, intentions, and behaviors. *Communication Monographs*, 83(4), 446-467.
- Calanchini, J., Moons, W. G., & Mackie, D. M. (2016). Angry expressions induce extensive processing of persuasive appeals. *Journal of Experimental Social Psychology*, 64, 88-98.
- Carpenter, C. J. (2015). A meta-analysis of the ELM's argument quality×processing type predictions. *Human Communication Research*, 41(4), 501-534.
- Chu, S. C., & Kamal, S. (2008). The effect of perceived blogger credibility and argument quality on message elaboration and brand attitudes: An exploratory study. *Journal of Interactive Advertising*, 8, 26-37.
- Cialdini, R. B. (2001). *Influence. Science and Practice* (4th Ed.). Boston, MA: Allyn & Bacon.
- Cohen, J. (1992). A power primer. *Psychological Bulletin*, 112(1), 155-159.
- De Graaf, A., Sanders, J., & Hoeken, H. (2016). Characteristics of narrative interventions and health effects: A review of the content, form, and context of narratives in health-related narrative persuasion research. *Review of Communication Research*, 4, 88-131.
- Dillard, J. P., Hunter, J. E., & Burgoon, M. (1984). Sequential-request persuasive strategies. *Human Communication Research*, 10(4), 461-488.
- Eerland, A., & Van den Bergh, H. (2016). Empirische basis van conclusies. *Tijdschrift voor Taalbeheersing*, 38(2), 139-146.
- Eisend, M. (2006). Two-sided advertising: A meta-analysis. *International Journal of Research in Marketing*, 23(2), 187-198.
- Eisend, M. (2011). How humor in advertising works: A meta-analytic test of alternative models. *Marketing Letters*, 22(2), 115-132.
- Eisend, M. (2017). Meta-Analysis in advertising research. *Journal of Advertising*, 46(1), 21-35.
- Eisend, M., & Tarrahi, F. (2016). The effectiveness of advertising: A meta-meta-analysis of advertising inputs and outcomes. *Journal of Advertising*, 45(4), 519-531.
- Feeley, T. H., Anker, A. E., & Aloe, A. M. (2012). The door-in-the-face persuasive message strategy: A meta-analysis of the first 35 years. *Communication Monographs*, 79(3), 316-343.
- Feeley, T. H., & Moon, S. I. (2009). A meta-analytic review of communication campaigns to promote organ donation. *Communication Reports*, 22(2), 63-73.
- Ferguson, C. J., & Heene, M. (2012). A vast graveyard of undead theories: Publication bias and psychological science's aversion to the null. *Perspectives on Psychological Science*, 7(6), 555-561.
- Fishbein, M., & Yzer, M. C. (2003). Using theory to design effective health behavior interventions. *Communication Theory*, 14(2), 164-183.
- Gollwitzer, P. M. (1999). Implementation intentions: Strong effects of simple plans. *American Psychologist*, 54(7), 493-503.
- Gollwitzer, P. M., & Sheeran, P. (2006). Implementation intentions and goal achievement: A meta-analysis of effects and processes. *Advances in Experimental Social Psychology*, 38, 69-119.

- Hoeken, H. (2001). Anecdotal, statistical, and causal evidence: Their perceived and actual persuasiveness. *Argumentation*, 15(4), 425-437.
- Jackson, S., & Jacobs, S. (1983). Generalizing about messages: Suggestions for design and analysis of experiments. *Human Communication Research*, 9, 169-181.
- Kahneman, D. (2011). *Thinking, Fast and Slow*. New York, NY: Farrar, Straus and Giroux.
- Kumkale, G. T., Albarracín, D., & Seignourel, P. J. (2010). The effects of source credibility in the presence or absence of prior attitudes: Implications for the design of persuasive communication campaigns. *Journal of Applied Social Psychology*, 40(6), 1325-1356.
- Lipsey, M. W., & Wilson, D. T. (2001). *Practical Meta-Analysis*. Thousand Oaks, CA: Sage.
- Manning, M. (2009). The effects of subjective norms on behaviour in the theory of planned behaviour: A meta-analysis. *British Journal of Social Psychology*, 48(4), 649-705.
- Meuffels, H. L. M., & Van den Bergh, H. (2006). De ene tekst is de andere niet: The language-as-a fixed-effect fallacy revisited: Statistische implicaties. *Tijdschrift voor Taalbeheersing*, 28(4), 323-345.
- Mulder, G. (2016). De kwaliteit van onderzoek. *Tijdschrift voor Taalbeheersing*, 38(2), 163-173.
- Noar, S. M., Hall, M. G., Francis, D. B., Ribisl, K. M., Pepper, J. K., & Brewer, N. T. (2016). Pictorial cigarette pack warnings: A meta-analysis of experimental studies. *Tobacco Control*, 25, 341-354.
- O'Keefe, D. J. (2002). The persuasive effects of variation in standpoint articulation. In F.H. van Eemeren (Ed.), *Advances in Pragma-Dialectics* (pp. 65-82). Amsterdam: Sic Sat.
- O'Keefe, D. J. (2015). Message generalizations that support evidence-based persuasive message design: Specifying the evidentiary requirements. *Health communication*, 30(2), 106-113.
- O'Keefe, D. J. (2017). Misunderstandings of effect sizes in message effects research. *Communication Methods and Measures*, 11(3), 210-219.
- O'Keefe, D. J., & Jensen, J. D. (2007). The relative persuasiveness of gain-framed and loss-framed messages for encouraging disease prevention behaviors: A meta-analytic review. *Journal of Health Communication*, 12, 623-644.
- O'Keefe, D. J., & Jensen, J. D. (2009). The relative persuasiveness of gain-framed and loss-framed messages for encouraging disease detection behaviors: A meta-analytic review. *Journal of Communication*, 59(2), 296-316.
- Petty, R. E., & Cacioppo, J. T. (1986). *Communication and Persuasion: Central and Peripheral Routes to Attitude Change*. New York, NY: Springer.
- Pham, M. T., Geuens, M., & De Pelsmacker, P. (2013). The influence of ad-evoked feelings on brand evaluations: Empirical generalizations from consumer responses to more than 1000 TV commercials. *International Journal of Research in Marketing*, 30(4), 383-394.
- Raju, S., Unnava, H. R., & Montgomery, N. V. (2009). The moderating effect of brand commitment on the evaluation of competitive brands. *Journal of Advertising*, 38(2), 21-36.
- Sheeran, P. (2002). Intention-behavior relations: A conceptual and empirical review. *European Review of Social Psychology*, 12(1), 1-36.
- Sheeran, P., Maki, A., Montanaro, E., Avishai-Yitshak, A., Bryan, A., Klein, W. M., Miles, E., & Rothman, A. J. (2016). The impact of changing attitudes, norms, and self-efficacy on health-related intentions and behavior: A meta-analysis. *Health Psychology*, 35(11), 1178-1188.
- Shen, F., Sheer, V. C., & Li, R. (2015). Impact of narratives on persuasion in health communication: A meta-analysis. *Journal of Advertising*, 44(2), 105-113.
- Sopory, P., & Dillard, J. P. (2002). The persuasive effects of metaphor: A meta-analysis. *Human Communication Research*, 28(3), 382-419.
- Steinmetz, H., Knappstein, M., Ajzen, I., Schmidt, P., & Kabst, R. (2016). How effective are behavior change interventions based on the theory of planned behavior? A three-level meta-analysis. *Zeitschrift für Psychologie*, 224(3), 216-233.

- Tannenbaum, M. B., Hepler, J., Zimmerman, R. S., Saul, L., Jacobs, S., Wilson, K., & Albarracín, D. (2015). Appealing to fear: A meta-analysis of fear appeal effectiveness and theories. *Psychological bulletin*, 141(6), 1178-1204.
- Tormala, Z.L., Briñol, P., & Petty, R.E. (2006). When credibility attacks: The reverse impact of source credibility on persuasion. *Journal of Experimental Social Psychology*, 42, 684-691.
- Webb, T. L., & Sheeran, P. (2006). Does changing behavioral intentions engender behavior change? A meta-analysis of the experimental evidence. *Psychological Bulletin*, 132(2), 249-268.
- Weerman, A., Hoeken, H., & Hornikx, J. M. A. (2016). Een kritische analyse van de manipulatie van argumentkwaliteit in reclameonderzoek. In D. van de Mierop, L. Buysse, & R. Coesemans, R.(Red.), *De macht van de taal: Taalbeheersingsonderzoek in Nederland en Vlaanderen* (pp. 309-321). Leuven: Acco.
- Wirtz, J. G., Sparks, J. V., & Zimbres, T. M. (2018). The effect of exposure to sexual appeals in advertisements on memory, attitude, and purchase intention: A meta-analytic review. *International Journal of Advertising*, 37(2), 168-198.
- Xu, Z., & Guo, H. (2018). Advantages of Anticipated Emotions over Anticipatory Emotions and Cognitions in Health Decisions: A Meta-Analysis. *Health Communication*, 33(5), 519-525.
- Zebregs, S., Van den Putte, B., Neijens, P., & De Graaf, A. (2015). The differential impact of statistical and narrative evidence on beliefs, attitude, and intention: A meta-analysis. *Health communication*, 30(3), 282-289.

Zijn concrete argumenten doorslaggevender?

Een experiment naar het effect van mate van concreetheid op sociale oordelen

Lettica Hustinx, Iris Hofstra en Anne Janssen
Radboud Universiteit Nijmegen

TVT 41 (1): 137–150

DOI: 10.5117/TVT2019.1.009.HUST

Abstract

Are concrete arguments more decisive than abstract arguments?

Research has shown that concrete language is more understandable, interesting and stored better in memory than abstract language (i.e. Sadoski, Goetz, & Rodriguez, 2000). Results are however inconclusive when it comes to the persuasive power of concrete language, which could be due to different operationalisations of the concept across studies. The current study aims to disentangle the concepts of concreteness and of details by manipulating the materials systematically. In an experiment, participants ($N = 207$) read a judicial case. On the basis of three pro arguments and three counter-arguments they were asked to judge a woman's fitness as a parent. When the pro arguments were formulated concretely, counter-arguments were formulated abstractly and vice versa. Concreteness comprised three versions: concrete, concrete with details and concrete with affective details. In order to test whether concrete arguments were more decisive (cf. *Availability Heuristic*), participants were also asked to recall the arguments. Results show a marginally significant effect of plain concreteness. However, when concrete arguments were accompanied by non-affective details, these arguments were significantly more convincing. In addition, concrete arguments were recalled better than abstract arguments. Findings are discussed with regard to the different operationalisations of concreteness across studies.

Keywords: concreteness, abstractness, details, persuasion, experiment

Inleiding

Een wijdverbreid schrijfadvis in onderwijs, voorlichting en journalistiek is om zoveel mogelijk concrete taal te gebruiken en abstract taalgebruik te vermijden, teneinde begrijpelijkheid en aantrekkelijkheid van teksten te verhogen. Onderzoek ondersteunt deze aanname: concreet taalgebruik blijkt inderdaad begrijpelijker, aantrekkelijker en beter reproduceerbaar te zijn dan abstracte taal (Douma, 1994; Sadoski, Goetz, & Rodriguez, 2000). De *Dual Coding Theory* (Paivio, 1986), die kortweg postuleert dat abstracte informatie enkel verbaal wordt verwerkt en opgeslagen terwijl concrete informatie ook nog een visuele representatie heeft, zou deze concreetheids-effecten kunnen verklaren.

Veel minder consensus is er over de vraag of concrete taal ook overtuigender is dan abstracte. De literatuur laat hier een wisselend beeld zien: soms blijkt concreet taalgebruik overtuigender en geloofwaardiger te zijn dan abstract taalgebruik (Hansen & Wänke, 2010; Reyes, Thompson, & Bower, 1980; Shedler & Manis, 1986), maar in veel gevallen komt dit ook niet naar voren (Frey & Eagly, 1993; Guadagno, Rhoads, & Sagarin, studie 2, 2011). Een van de verklaringen zou kunnen zijn dat concreetheid in de studies op veel verschillende manieren geoperationaliseerd wordt. Dit kan variëren van het vervangen van ‘statische’ werkwoorden door actiewerkwoorden in stellingen (Hansen & Wänke, 2010), tot het toevoegen van uitgebreide details (o.a. Bensi, Nori, Strazzari, & Giusberti, 2003; Reyes et al., 1980; Shedler & Manis, 1986), die al dan niet relevant zijn voor de (argumentatieve) inhoud van de tekst (Guadagno et al., 2011).

In dit artikel proberen we een overzicht te geven van de verschillende concreetheidsonderzoeken en hun concreetheidmanipulaties, en proberen we in een persuasie-experiment concreetheid systematischer te onderzoeken. In onze zoektocht naar concreetheid zijn er grote raakvlakken met het diffusere concept levendigheid, waar concreetheid de talige realisering van vormt (Nisbett & Ross, 1980; Taylor & Thompson, 1982). In dit artikel worden de levendighedsstudies, voor zover deze concreetheid en details als manipulatie gebruiken, ook besproken. Achtereenvolgens komen aan de orde: onderzoeken naar oordelen over waarheid en juridische cases, de rol van details in persuasie-onderzoek, en de mate waarin die details een emotie oproepen. De inleiding wordt afgesloten met een bespreking van recente studies die intersubjectief verkregen scores van concreetheid leveren.

Studies naar de persuasieve invloed van concreetheid, levendigheid en details op oordelen spelen zich meestal af in een setting waarin participanten de perceptie van waarheid van stellingen moeten inschatten (Hansen

& Wänke, 2010) of een juridische of maatschappelijke casus moeten beoordelen (Reyes et al., 1980; Shedler & Manis, 1986). In een experiment naar de waarheidsperceptie, waarin ware en onware stellingen werden aangeboden die concreet ofwel abstract geformuleerd werden, bleken de concreet geformuleerde stellingen als meer 'waar' te worden beoordeeld dan de abstracte, ongeacht hun werkelijke waarheid. Concreetheid was in dit onderzoek volgens het *Linguistic Category Model* geoperationaliseerd als actiewerkwoorden (bijvoorbeeld 'schrijft') die afgezet werden tegen dezelfde stellingen met de meer abstracte statische werkwoorden (bijvoorbeeld 'is' auteur van) (Hansen & Wänke, 2010).

In de studies met een juridische setting (bijvoorbeeld Reyes et al., 1980; Shedler & Manis, 1986) worden participanten zowel pro- als contra-argumenten voorgelegd, waarvan de helft concreet en levendig gemaakt is en de andere helft minder levendig door weglaten van details. De bevindingen laten zien dat levendige en gedetailleerde argumenten, zowel aan de verdedigende als de beschuldigende kant, zwaarder wegen bij oordelen dan de minder levendige argumenten. De stijl van het argument draagt dus bij aan de doorslaggevende kracht van het argument.

De vraag hierbij is telkens of de Availability Heuristic (Tversky & Kahneman, 1973) de resultaten kan verklaren. Deze heuristiek die voortbouwt op de Dual Coding Theory (Paivio, 1986), stelt dat concrete en levendige informatie vanwege de veronderstelde voorstelbaarheid gemakkelijker op te halen is uit het geheugen, en derhalve meer invloed heeft op oordeeltaken dan niet-concrete taal die minder makkelijk op te halen is. De concrete informatie zou dus bij een reproductietaak dan ook meer gereproduceerd moeten worden, iets wat ook regelmatig wordt gevonden (Reyes et al., 1980; Shedler & Manis, 1986).

Het onderhavige onderzoek is een gedeeltelijke replicatie van het onderzoek van Shedler en Manis (1986). In hun studie moesten participanten zich inbeelden lid te zijn van een burgerjury om vervolgens te oordelen over de geschiktheid van mevrouw Johnson als moeder voor haar zoon. Dit oordeel werd gevormd op basis van zestien argumenten, die elk een levendige en een niet-levendige versie hadden. Dit leidde tot twee tekstcondities: acht levendige (concrete) pro-argumenten die in combinatie met acht niet-levendige contra-argumenten werden aangeboden, en in de andere tekstconditie acht niet-levendige pro-argumenten in combinatie met concrete contra-argumenten. Levendigheid werd gemanipuleerd door toevoeging van concrete en kleurrijke details die voor het oordeel niet relevant waren. Ter illustratie, een pro-argument in de niet-verlevendigde vorm luidde: *'Mrs. Johnson sees to it that her child washes and brushes his teeth before bed*

time.' Aan de verlevendigde versie werd daaraan het volgende detail toegevoegd: *'He uses a Star Wars toothbrush that looks like Darth Vader.'* (p. 27).

De resultaten lieten zien dat het verwachte levendigheidseffect werd gevonden: verlevendigde pro-argumenten in combinatie met niet-levendige contra-argumenten leverden een positiever geschiktheidsoordeel op dan verlevendigde contra-argumenten in combinatie met niet-levendige pro-argumenten. Ook werden levendige argumenten beter gereproduceerd. Een regressieanalyse liet echter zien dat het resultaat niet verklaard kon worden door een betere geheugenopslag en dus ook niet door de Availability Heuristic, een verklaring die wél in de studie van Reyes et al. (1980) werd geopperd.

In levendigheidsonderzoek zijn de gepresenteerde toegevoegde details soms irrelevant, zoals de details in het voorbeeld hierboven. Toch zijn er ook studies die aantonen dat het toevoegen van irrelevante details afleidt van de centrale boodschap en daardoor het persuasieve effect juist ondermijnt (Guadagno et al., 2011; Kaminski, Sloutsky, & Heckler, 2013). Relevante details daarentegen zouden wel de overtuigingskracht vergroten (Guadagno et al., 2011; Pettus & Diener, 1977).

In de literatuur wordt in enkele gevallen een link gelegd tussen concreetheid en emotionele respons. En inderdaad, bij bestudering van het materiaal van verschillende concreetheids- en levendigheidsonderzoeken, lijkt er niet alleen sprake te zijn van (ir)relevantie van de details, maar lijkt ook de mate waarin een detail een affectieve respons oproept een rol te spelen: zie bijvoorbeeld het detail hierboven over de tandenborstel. De affectie die dat oproept zou deels de resultaten van het oordeel kunnen verklaren.

Er is nog weinig onderzoek gedaan naar emotionaliteit als onderliggende variabele in concreetheids- en levendigheidsonderzoek. Wel is er een hoge correlatie tussen voorstelbaarheid (*imagery*) en emotionele respons gevonden (Goetz, Sadoski, Stowe, Fetsco, & Kemp, 1993). Daarnaast werd in een concreetheidsonderzoek aannemelijk gemaakt dat affectiviteit, meer dan concreetheid, leidt tot een snellere reactietijd en grotere nauwkeurigheid in woordherkenning (Kousta, Vigliocco, Vinson, Andrews, & Del Campo, 2011). In onderzoek van Pettus en Diener (1977) bleek dat emotie en details elkaar versterken. Daarom vragen we ons af of affectiviteit in details ook een rol zou kunnen spelen bij overtuigingskracht.

Recent onderzoek maakt het mogelijk om de mate van concreetheid op woordniveau vast te stellen, iets wat ten tijde van de eerder besproken studies nog niet beschikbaar was. Zo zijn er concreetheidsscores van 30.000 Nederlandstalige woorden verzameld (Brysbaert, Stevens, De Deyne, Voorspoels, & Storms, 2014), door proefpersonen op een vijfpuntsschaal een oordeel te laten geven over de mate van concreetheid per woord. Ook

de studie naar determinanten van het concept concreetheid (Spooren, Hustinx, Aben, & Turkenburg, 2015) biedt houvast: vooral zintuiglijke waarneembaarheid voorspelt concreetheid, gevolgd door specificiteit en filmbaarheid. Dankzij deze metingen kan concreetheidsonderzoek met objectievere maatstaven uitgevoerd worden.

Samenvattend roept de literatuur over concreetheid en levendigheid de vragen op waardoor het concreetheidseffect in oordelexperimenten wordt veroorzaakt, wat de rol van details daarbij is, en in hoeverre affectiviteit een rol zou kunnen spelen. De onderzoeksvraag in deze studie is: In hoeverre heeft de mate van concreetheid van argumenten invloed op een sociaal oordeel, wat is de rol van details en in hoeverre speelt emotionaliteit in details een rol?

In een experiment worden de drie stijlvarianten ‘concreet’, ‘concreet met detail’ en ‘concreet met affectief detail’ verwerkt in argumenten en in één tekst gecombineerd aangeboden met abstract geformuleerde argumenten. De verwachting is dat concrete argumenten zwaarder wegen dan abstracte argumenten (H₁) en dat concrete argumenten met details zwaarder wegen in de oordeelsvorming dan abstracte argumenten (H₂). Verder verwachten we intuïtief dat de variant met de concrete argumenten met affectieve details ook een concreetheidseffect zal laten zien, ondanks dat de literatuur niet direct uitsluitsel geeft over de rol van affectiviteit.

Het gaat in dit experiment dus niet om een stijlmanipulatie per tekstconditie – een conditie heeft immers zowel concrete als abstracte argumenten – maar om de relatieve kracht van de concreetheidsvariant die in contrast met abstractheid wordt aangeboden (in navolging van Reyes et al., 1980; Shedler & Manis, 1986). De ene tekstconditie met concrete proargumenten en abstracte contra-argumenten wordt dus vergeleken met de andere tekstconditie waarin de contra-argumenten concreet en de proargumenten abstract zijn.

Tot slot verwachten we dat concrete argumenten beter gereproduceerd worden dan abstracte (H₃), waardoor de voorspellingen van de Availability Heuristic zouden worden ondersteund.

Methode

Om de onderzoeksvragen te beantwoorden is een experiment opgezet waarin participanten zich moesten inbeelden lid te zijn van een burgerjury, en moesten oordelen over de vraag hoe geschikt moeder Ellen was om te zorgen voor haar zoon Tim.

Materiaal en design

De participanten werd eerst een casus aangeboden over een alleenstaande moeder die voor haar zevenjarige zoon zorgt. Er werd vermeld dat de moeder op sommige onderdelen van de verzorging goed maar op andere slecht presteerde en dat Kinderbescherming overwoog om haar zoon uit huis te plaatsen. Participanten werd gevraagd om op basis van argumenten te oordelen over de geschiktheid van de moeder. Er zijn zes tekstcondities geconstrueerd die allemaal drie pro- en drie contra-argumenten bevatten. Als de pro-argumenten concreet waren, werden de contra-argumenten in hun abstracte variant gepresenteerd. Daarbij zijn aan de concrete argumenten (pro en contra) details in variant 2 en affectieve details in variant 3 toegevoegd. In Tabel 1 wordt het een en ander verduidelijkt.

Tabel 1 Design met de 3 concreetheidsvarianten

Variant	Tekstconditie	Concreet	Abstract
1	1	Pro-argumenten concreet	Contra-argumenten abstract
	2	Contra-argumenten concreet	Pro-argumenten abstract
2	3	Pro-argumenten concreet met details	Contra-argumenten abstract
	4	Contra-argumenten concreet met details	Pro-argumenten abstract
3	5	Pro-argumenten concreet met affectieve details	Contra-argumenten abstract
	6	Contra-argumenten concreet met affectieve details	Pro-argumenten abstract

Bij de keuze voor welke argumenten moesten pleiten voor en tegen de geschiktheid van de moeder is uitgegaan van een vooronderzoek naar het belang van verschillende onderdelen van de opvoeding. Op basis hiervan werden de voorargumenten 'goede voeding', 'op tijd naar bed' en 'goede lichamelijke verzorging' gekoppeld aan de tegenargumenten 'geen problemen kunnen vertellen', 'onhygiënisch huis' en 'te laat komen op school', die paarsgewijs even zwaar wogen. Van elk argument (drie pro en drie contra) zijn vier stijlmanipulaties gemaakt: abstract, concreet, concreet met detail, concreet met affectief detail. Pro- en contra-argumenten werden alternerend aangeboden, waarmee dus ook de stijl van de argumenten afwisselde.

Concreetheid werd op een systematische manier gemanipuleerd. Inhoudswoorden en frequentiewoorden werden geselecteerd en gescoord op concreetheid, met als eis dat concrete argumenten gemiddeld een punt hoger scoorden op de concreetheidsschaal van Brysbaert et al. (2014) dan de abstracte argumenten. De gemiddelde concreetheidsscore van de concrete argumenten (concreet, concreet detail, concreet affectief detail) is 3.44 ($SD = 0.27$) en die van de abstracte argumenten 2.39 ($SD = 0.26$). Verder

is ervoor gezorgd dat werkwoorden met name filmisch en zelfstandig naamwoorden zo specifiek mogelijk waren (vergelijk Spooren et al., 2015).

De toegevoegde details gingen om zaken die passen bij de inhoud van het argument, in lijn met de relevantie-eisen van Guadagno et al. (2011). De affectieve details riepen een emotie op die congruent was met het argument (vergelijk Smith & Shaffer, 2000). De concrete en abstracte pro- en contra-argumenten zonder details waren in aantal tekens onderling even lang en de toegevoegde details ook, om lengte-effecten uit te sluiten.

Een pilotstudie wees uit dat details als detail werden ervaren en niet als extra argument, en ook dat affectieve details daadwerkelijk meer emotie opwekten dan gewone details. Tabel 2 geeft een overzicht van de vier stijlmanipulaties van een pro-argument.

Tabel 2 Voorbeelden van het materiaal (aantal tekens tussen haakjes) in de vier stijlmanipulaties

Concreet pro	Concreet pro met detail	Concreet pro met affectief detail	Abstract pro
Ellen vindt het belangrijk dat Tim gezond eet bij ontbijt, lunch en avondeten. Vier keer per week kookt ze groenten, vlees en aardappelen voor Tim. (147)	[concreet pro-argument] + Wanneer ze kipfilet eten, geeft Ellen Tim daar ook altijd wat tomatenketchup bij. (229)	[concreet pro-argument] + Wanneer ze kipfilet eten, maakt ze daar voor Tim een gezichtje van ketchup op. (226)	Ellen vindt evenwichtige voeding belangrijk voor Tim. Zo maakt ze regelmatig een maaltijd voor hem waarin de belangrijkste voedingsstoffen zitten. (146)

NB: De (affectieve) details zijn toegevoegd aan concrete argumenten, hetgeen het hogere aantal tekens verklaart.

Instrumentatie en procedure

De geschiktheid van de moeder werd bevraagd met vier elf-puntschalen, die deels zijn overgenomen uit Shedler en Manis (1986). De vragen luiden als volgt: (1) Hoe geschikt vind je Ellen als moeder voor Tim? (2) In welke mate vind je de situatie waarin Tim zich bevindt problematisch? (3) In welke mate vind je dat Ellen hulp nodig heeft bij de opvoeding van Tim? (4) Stel dat de situatie van Ellen en Tim niet verandert en er komt geen hulp. In welke mate ben je het dan eens met de stelling: 'Tim moet uit huis geplaatst worden'? Een score van 1 op de schaal was ontkennend, 11 was bevestigend. De betrouwbaarheid van de vier vragen was adequaat ($\alpha = 0.73$). Na vragen over personalia kregen de participanten een reproductietaak waarbij hen gevraagd werd om, zonder terug te kijken, alles van de argumenten op te schrijven wat zij zich nog konden herinneren.¹

Participanten

In totaal hebben 223 studenten van 17 tot 30 jaar ($M = 19.32$) van de Radboud Universiteit en de Hogeschool van Arnhem en Nijmegen de vragenlijst ingevuld. Na uitsluiting van niet complete vragenlijsten zijn de antwoorden van 207 participanten geanalyseerd. Het aantal vrouwen ($N = 139$) lag hoger dan het aantal mannen ($N = 68$). De man-vrouwverhouding was gelijk verdeeld over de tekstcondities.

Resultaten

Geschiktheidsoordelen

Om de resultaten van de oordelen te analyseren zijn *Welch two sample t-tests* gedaan tussen de tekstcondities met dezelfde concreetheidsvarianten. Alle verschillen die op basis van gerichte hypothesen zijn uitgevoerd, zijn eenzijdig getoetst.

Uit een t-toets voor de twee tekstcondities waarin concrete (ongedetailleerde) argumenten worden afgezet tegen hun abstracte pendanten (variant 1, zie Tabel 3) blijkt een marginaal verschil tussen het oordeel op basis van tekstconditie 1 en tekstconditie 2 op te treden ($t(66.00) = 1.48, p = .072$): participanten die de argumenten in tekstconditie 1 (concreet-pro in combinatie met abstract-contra) hebben gelezen beoordelen de moeder als iets geschikter ($M = 6.15$) dan participanten die 'de tegenhanger' in tekstconditie 2 (concreet-contra in combinatie met abstract-pro) lazen ($M = 5.59$). Dit verschil is echter een trend.

Tabel 3 Gemiddelde geschiktheidsoordelen van de vier vragen tezamen als functie van stijlvarianten (SD tussen haakjes)

	Pro concreet Contra abstract	Contra concreet Pro abstract
Variant 1: Concreet (1 versus 2)	6.15 (1.61)	5.59 (1.51) [†]
Variant 2: Concreet detail (3 versus 4)	6.25 (1.55)	5.27 (1.46)**
Variant 3: Concreet affectief detail (5 versus 6)	5.94 (1.14)	5.75 (1.37)

NB 1: Verschillen met superscript [†] = marginaal significant; ** = significant $p < .01$.

NB 2: De vragen 2, 3, en 4 zijn omgepoold zodat hogere gemiddelden wijzen op positievere oordelen.

In concreetheidsvariant 2, waarin de oordelen op basis van tekstcondities 3 en 4 (met de toegevoegde details) met elkaar worden vergeleken verschillen de oordelen significant van elkaar ($t(65.07) = 2.67, p = .009$). Variant 3, waarin tekstcondities 5 en 6 tegen elkaar worden afgezet (met de affectieve details) laat geen verschil zien tussen beide geschiktheidsoordelen ($t(68.30) = 0.64, p = .524$).

Recall

Om na te gaan of de Availability Heuristic (Tversky & Kahneman, 1973) de resultaten zou kunnen verklaren, zijn de argumenten opgedeeld in *idea-units*. Bijvoorbeeld het argument ‘Ellen vindt het belangrijk dat Tim gezond eet bij ontbijt, lunch en avondeten. / Vier keer per week / kookt ze groenten, vlees en aardappelen voor Tim’ bevat drie idea-units, die hier onderstreept zijn. Het scoren ging als volgt: als de inhoud van de idea-unit geparafraseerd of gedeeltelijk goed was leverde dat een score van één punt op, wanneer het gereproduceerde vrijwel identiek was aan het materiaal, was de score twee punten. In totaal waren er maximaal 22 punten te behalen (elf idea-units) zowel bij de concrete als de abstracte pro-argumenten, en 20 punten (tien idea-units) voor de contra-argumenten. De details leverden in beide varianten maximaal tien (pro-argumenten) respectievelijk twaalf (contra-argumenten) punten op.

Omdat in concreetheidsvarianten 2 en 3 (die met de details) de concrete informatie langer was dan de abstracte zijn er per variant eerst analyses uitgevoerd op de percentages gereproduceerde idea-units. In Tabel 4 staan de gemiddelde percentages gereproduceerde idea-units, voor alle drie de concreetheidsvarianten. Hier zijn de gegevens van de concrete *recall* (zowel van de pro-argumenten als de contra-argumenten) afgezet tegen de abstracte recall (ook van de pro- en contra-argumenten). Er zijn drie Welch two sample t-tests gedaan om te kijken of bij elke variant de concrete informatie beter werd gereproduceerd dan de abstracte.

De concrete informatie zonder details wordt beter gereproduceerd dan de abstracte ($t(105.15) = 4.84, p < .001$). Dit geldt ook voor de concrete informatie met details ($t(133.37) = 1.99, p = .049$ en voor de affectieve details ($t(138.79) = 3.42, p < .001$). De concrete argumenten worden in al hun varianten dus meer en beter gereproduceerd dan de abstracte argumenten.

Tabel 4 Gemiddeld aantal gereproduceerde idea-units van concrete en abstracte argumenten in percentages (SD tussen haakjes)

	Concreet (pro en contra)	Abstract (pro en contra)
Variant 1: Concreet (1 versus 2)	36.2 (19.6)	23.1 (10.9)**
Variant 2: Concreet detail (3 versus 4)	38.5 (18.5)	32.0 (19.8)*
Variant 3: Concreet affectief detail (5 versus 6)	35.8 (19.9)	24.9 (18.1)**

NB 1: * = $p < .05$, ** = $p < .001$

Als we kijken naar de verschillen in absolute aantallen tussen concreet en abstract waarbij alleen het concrete deel wordt geteld, zonder de details

per variant, dan zien we wederom de reproductieverschillen in elke variant optreden. Bij de concrete versus de abstracte argumenten vinden we uiteraard weer het verschil tussen de concrete ($M = 7.62$, $SD = 4.09$) en abstracte argumenten ($M = 4.81$, $SD = 2.26$) ($t(104.53) = 4.96$, $p < .001$). Bij variant 2 zien we dit verschil ook tussen de concrete argumenten met details ($M = 8.72$, $SD = 4.32$) versus de abstracte ($M = 6.66$, $SD = 3.93$) ($t(132.81) = 2.91$, $p = .004$). Ook bij de derde variant is het verschil tussen de concrete argumenten met affectieve details ($M = 7.39$, $SD = 4.27$) en hun abstracte tegenhangers ($M = 5.18$, $SD = 3.70$) significant ($t(137.21) = 3.3$, $p = .001$). Samenvattend worden concrete argumenten, zowel absoluut als relatief, beter gereproduceerd dan abstracte argumenten.

Discussie

In ons experiment oordeelden participanten aan de hand van pro- en contra-argumenten over de geschiktheid van een moeder voor haar zoon. Deze argumenten waren, anders dan in eerder onderzoek (Pettus & Diener, 1977; Reyes et al., 1980; Shedler & Manis, 1986), systematisch gemanipuleerd op concreetheid en gedetailleerdheid om op die manier na te gaan welke realisatie van concreetheid overtuigingseffecten tot gevolg kan hebben. Hadden we dit onderscheid niet gemaakt, dan was onze uitkomst vergelijkbaar geweest met de voorgaande studies. Een achteraf-analyse laat namelijk zien dat als concreetheid in alle stijlvarianten ('op één hoop gegooid') tezamen afgezet wordt tegen de abstracte pendant er een duidelijk effect optreedt: concrete pro-argumenten in combinatie met abstracte tegenargumenten geven een positiever oordeel ($M = 6.11$, $SD = 1.44$), dan concrete tegenargumenten in combinatie met abstracte pro-argumenten ($M = 5.54$, $SD = 1.45$), ($F(1, 204) = 7.87$, $p = .006$). Wij hebben geprobeerd om op basis van wat we in de literatuur observeerden drie mogelijke concretheidsvarianten te onderscheiden, en dan zien we dat het wat ingewikkelder ligt.

Kijkend naar de geschiktheidsoordelen over de argumenten van de eerste variant, zien we dat concrete pro-argumenten met abstracte contra-argumenten voor een marginaal positiever geschiktheidsoordeel zorgen dan concrete contra-argumenten met abstracte pro-argumenten. Hypothese 1 kan dus niet bevestigd worden; er is hooguit sprake van een trend. Als je concreetheid heel nauwkeurig manipuleert volgens de recent geleverde scores van Brysbaert et al. (2014), die uiteraard niet ideaal maar toch een betere indicatie van concreetheid geven dan intuïtief ontworpen experimenteel materiaal, dan lijkt het concretheidseffect niet overtuigend op te treden.

Wanneer er niet-affectieve details worden toegevoegd aan de concrete argumenten, zoals in variant twee, hebben deze wel significant meer invloed op het geschiktheidsoordeel dan abstracte argumenten: concrete pro-argumenten met details in combinatie met abstracte contra-argumenten zorgen voor een significant positiever geschiktheidsoordeel dan concrete contra-argumenten met details in combinatie met abstracte pro-argumenten. Hypothese 2 is daarmee bevestigd. Dit resultaat is in lijn met de bevindingen van de studies van Reyes et al. (1980) en Shedler en Manis (1986), waarin het materiaal ook door middel van details werd gemanipuleerd. Blijkbaar is er meer dan enkel een linguïstische realisering (vergelijk Brysbaert et al., 2014; Spoores et al., 2015) nodig om concreetheidseffecten te kunnen bewerkstellingen: details lijken een belangrijke bijdrage te leveren. Dit is eigenlijk opmerkelijk, aangezien abstracte argumenten in ons onderzoek ('belangrijkste voedingsstoffen') inhoudelijk meer omvatten dan de concrete invulling daarvan ('groenten, vlees en aardappelen'), laat staan dat details er argumentatief toe doen ('tomatenketchup').

Een punt van kritiek zou kunnen zijn dat de gevonden effecten verklaard kunnen worden door lengte. Argumenten met details zijn overtuigender dan hun abstracte tegenhangers, maar ook langer. Een univariate variantieanalyse, waarin tekstconditie 1 en 2 'zonder details' werden afgezet tegen tekstconditie 3 en 4 'met details' laat zien dat lengte geen covariaat is (Wilks Lambda, $F(4, 128) = 1.31, p = .272$). En als tekstconditie 1 en 2 'zonder details' afgezet worden tegen tekstconditie 5 en 6 'met affectieve details' blijkt lengte ook geen invloed te hebben (Wilks Lambda, $F(4, 132) = 0.34, p = .840$).

Onze resultaten lijken het idee te rechtvaardigen dat concreetheid en details te onderscheiden concepten zijn. Voorwaarde is volgens de literatuur wel dat de details relevant zijn (Guadagno et al., 2011) en congruent (Smith & Shaffer, 2000) met de inhoud van de argumenten. In ons experiment hebben we gepoogd de details daarmee in overeenstemming te ontwerpen. Of we daarin geslaagd zijn valt nog te bezien. De argumenten met affectieve details, de derde variant, leggen in tegenstelling tot onze intuïtie blijkbaar geen zwaarder gewicht in de schaal dan hun abstracte tegenhangers. Het is niet duidelijk waarom deze details niet het persuasieve voordeel opleveren zoals de 'gewone details' dat doen. Affectiviteit vormt volgens Nisbett en Ross (1980) een essentieel onderdeel van levendigheid, en zou daarom een extra duit in het zakje kunnen doen. Wellicht zijn de affectieve details, zoals 'een gezichtje van ketchup maken' of 'de rode badmat is versleten en Tim wil er niet met zijn blote voetjes op staan' toch net wat te irrelevant. We hebben in de voorstudie gevonden dat er meer emotie werd

opgeroepen door de affectieve details dan door de ‘gewone’ details, maar we hebben niet naar relevantie gevraagd en dat zou in vervolgonderzoek echt aan te bevelen zijn.

Om te achterhalen of de Availability Heuristic (Tversky & Kahneman, 1973) de gevonden resultaten kan verklaren, is een reproductietaak uitgevoerd. Concrete argumenten werden, ook zonder de details, zowel proportioneel als absoluut, meer gereproduceerd dan abstracte argumenten, waarmee H3 wordt ondersteund. Dit wijst op betere beschikbaarheid en zo kunnen concrete argumenten meer invloed gehad hebben op de oordelen. Toch komen de oordeelresultaten in de verschillende varianten niet overeen met die van de recall. De Availability Heuristic kan daarom de resultaten niet verklaren. Blijkbaar spelen er nog andere onderliggende factoren een rol, net zoals Shedler en Manis (1986) suggereren. In elk geval zijn deze bevindingen wel in lijn met het frequent gevonden recall-effect, hetgeen de Dual Coding Theory ondersteunt.

Een mogelijke verbetering bij vervolgonderzoek zou kunnen zijn om, net als bij Reyes et al. (1980) en Shedler en Manis (1986), dezelfde onderwerpen in de opvoeding (bijvoorbeeld ‘bedtijd’) zowel als pro- en als contra-argument op te nemen. Nadeel is dat een onderwerp als bijvoorbeeld ‘voeding’ dan niet gebruikt kan worden omdat, zoals bleek uit ons vooronderzoek, ‘je kind slecht te eten geven’ relatief zwaarder meetelt bij een geschiktheidsoordeel dan ‘je kind goed te eten geven’.

Een andere suggestie voor vervolgonderzoek is om ook concrete pro-argumenten af te zetten tegen concrete contra-argumenten, en dit ook te doen voor de abstracte pro- en contra-argumenten. Zo kun je beter onderscheid maken tussen wat nu de bijdrage van de stijl is en wat de kracht van de argumenten op hun beurt doet. Want ook al zijn de thema’s van de argumenten in een pre-test getoetst op hun belang in de opvoeding van kinderen, het kan nog steeds zo zijn dat de kracht van de pro-argumenten niet gelijk is aan die van de contra-argumenten, en dat alleen in concreetheidsvariant 2 het krachtsverschil wordt overstemd door de concreetheidsvariant.

Ons onderzoek laat het belang zien van recentelijk geleverde mogelijkheden om concreetheid op een objectievere manier in experimenteel materiaal te manipuleren. Ook hebben we getracht aannemelijk te maken dat bij toekomstig concreetheidsonderzoek de begrippen levendigheid, concreetheid en gedetailleerdheid duidelijk begrensde onderwerpen zouden moeten zijn binnen het taalbeheersingsonderzoek.

Noot

1. Ook begrijpelijkheid en voorstelbaarheid zijn bevraagd, maar deze vragen zijn niet goed geïnterpreteerd door de participanten. Ook is de participanten gevraagd om in de door hun gereproduceerde argumenten te onderstrepen wat de doorslag gaf bij hun oordeel, maar deze gegevens worden niet gerapporteerd.

Bibliografie

- Bensi, L., Nori, R., Strazzari, E., & Giusberti, F. (2003). Vividness in judgements of guilt. *Perceptual and Motor Skills*, 97(3), 1133-1136.
- Brybaert, M., Stevens, M., De Deyne S., Voorspoels, W., & Storms, G. (2014). Norms of age of acquisition and concreteness for 30,000 Dutch words. *Acta Psychologica*, 150, 80-84.
- Douma, P. (1994). Wees zo concreet mogelijk. Schrijfadvisers over concreet en abstract taalgebruik. *Tijdschrift voor Taalbeheersing*, 16(1), 16-31.
- Frey, K.P., & Eagly, A.H. (1993). Vividness can undermine the persuasiveness of messages. *Journal of Personality and Social Psychology*, 65(1), 32-44.
- Goetz, E.T., Sadoski, M., Stowe, M.L., Fetsco, T.G., & Kemp, S.G. (1993). Imagery and emotional response in reading literary text: Quantitative and qualitative analyses. *Poetics*, 22(1-2), 35-49.
- Guadagno, R.E., Rhoads, K.V.L., & Sagarin, B.J. (2011). Figural vividness and persuasion capturing the "elusive" vividness effect. *Personality and Social Psychology Bulletin*, 37(5), 626-639.
- Hansen, J., & Wänke, M. (2010). Truth from language and truth from fit: The impact of linguistic concreteness and level of construal on subjective truth. *Personality and Social Psychology Bulletin*, 36(11), 1576-1588.
- Kaminski, J.A., Sloutsky, V.M., & Heckler, A.F. (2013). The cost of concreteness: The effect of nonessential information on analogical transfer. *Journal of Experimental Psychology: Applied*, 19(1), 14-29.
- Kousta, S.T., Vigliocco, G., Vinson, D.P., Andrews, M., & Del Campo, E. (2011). The representation of abstract words: Why emotion matters. *Journal of Experimental Psychology: General*, 140(1), 14-34.
- Nisbett, R., & Ross, L. (1980). *Human Inference: Strategies and Shortcomings of Social Judgment*. Englewood Cliffs, NJ: Prentice-Hall.
- Paivio, A. (1986). *Mental Representations: A Dual Coding Approach*. Oxford: Oxford University Press.
- Pettus, C., & Diener, E. (1977). Factors affecting the effectiveness of abstract versus concrete information. *The Journal of Social Psychology*, 103(2), 233-242.
- Reyes, R.M., Thompson, W.C., & Bower, G.H. (1980). Judgmental biases resulting from differing availabilities of arguments. *Journal of Personality and Social Psychology*, 39(1), 2-12.
- Sadoski, M., Goetz, E.T., & Rodriguez, M. (2000). Engaging texts: Effects of concreteness on comprehensibility, interest, and recall in four text types. *Journal of Educational Psychology*, 92(1), 85-95.
- Shedler, J., & Manis, M. (1986). Can the availability heuristic explain vividness effects? *Journal of Personality and Social Psychology*, 51(1), 26-36.
- Smith, S.M., & Shaffer, D.R. (2000). Vividness can undermine or enhance message processing: The moderating role of vividness congruency. *Personality and Social Psychology Bulletin*, 26(7), 769-779.

- Spooren, W.P.M.S., Hustinx, L.G.M.M., Aben, J., & Turkenburg, E. (2015). Concreetheid onder de loep. In M. Boogaard, B. van den Bogaerde, S. Bacchini, M. Curcic, N. de Jong, E. le Pichon & L. Rasier (Red.), *Artikelen van de achtste Anéla Conferentie Toegepaste Taalwetenschap 2015* (pp. 97-110). Delft: Eburon.
- Taylor, S.E., & Thompson, S.C. (1982). Stalking the elusive "vividness" effect. *Psychological Review*, 89(2), 155-181.
- Tversky, A., & Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. *Cognitive Psychology*, 5(2), 207-232.

De invloed van het standpunt op de beoordeling van *ad populum*-argumentatie

Henrike Jansen
Universiteit Leiden

TVT 41 (1): 151–165

DOI: 10.5117/TVT2019.1.010.JANS

Abstract

The influence of the standpoint on the evaluation of ad populum argumentation

This paper presents guidelines for the analysis and evaluation of *ad populum* argumentation. For both of these tasks the type of standpoint is considered as a relevant factor. The conclusions are based on a critical examination of the argumentation theoretical literature regarding this type of argument and in completing and systematizing the insights drawn from it. *Ad populum*-argumentation supporting a descriptive standpoint can only be reasonable if the appeal to the opinion of a lot of people is supplemented with an extra coordinatively linked argument. *Ad populum* argumentation supporting a prescriptive standpoint should be treated according to its function in two contexts in which it typically occurs: advertising and politics. An analysis of the first context shows that this kind of *ad populum* actually consists of complex argumentation in which the appeal to the people supports an evaluative standpoint. This kind of *ad populum* is assessed as unreasonable, albeit for different reasons than provided in the literature. *Ad populum* argumentation in the context of political deliberation – ‘procedural *ad populum* argumentation’ – is regarded as argumentation that is not inherently fallacious, but still too weak to function as the only support for a policy standpoint.

Keywords: *ad populum*, democratic decision rule, ethical and aesthetical criteria, shared starting points, advertising, deliberation, type of standpoint

Inleiding

Met *ad populum*-argumentatie wordt betoogd dat een standpunt geaccepteerd moet worden omdat een aanzienlijke groep mensen het ermee eens is.¹ Traditioneel wordt een dergelijke redenering als een drogreden beschouwd; de geschiedenis laat immers talloze voorbeelden zien van wijdverbreide maar onjuiste opvattingen, zoals het geloof in een platte aarde. In dit licht zou het enerzijds als een paradoxale ontwikkeling kunnen worden beschouwd dat er steeds meer politici zijn die een beroep doen op de ‘wil van het volk’ om zo het beleid dat zij voorstaan meer gewicht te geven. Anderzijds lijkt het echter overdreven om dergelijke *ad populum*-argumenten als drogredenen te beschouwen: zeggen dat bepaald beleid moet worden geaccepteerd omdat veel mensen er voorstander van zijn is van een andere orde dan zeggen dat iets het geval is omdat veel mensen denken dat het zo is.

In deze bijdrage onderzoek ik de hypothese dat bij het bepalen van de redelijkheid van *ad populum*-argumentatie het standpunt een rol speelt. In de pragma-dialectische argumentatietheorie wordt een onderscheid gemaakt tussen drie soorten standpunten (Van Eemeren, 2010, pp. 1-2). In een descriptief standpunt geeft de propositie een stand van zaken aan, in heden, verleden of toekomst. Dergelijke standpunten hebben als basisvorm: ‘X is het geval’. In een evaluerend standpunt geeft de propositie een waardeoordeel aan: er wordt beweerd dat iets goed of slecht, mooi of lelijk, enzovoort is. De basisvorm van dit type standpunt is: ‘X is [waardeoordeel]’. Een prescriptief standpunt bevat een propositie waarin de wenselijkheid van het doen of nalaten van een bepaalde handeling wordt uitgedrukt, met als basisvorm: ‘X moet (niet) worden uitgevoerd’. Een ander type standpunt kan van een argument een drogreden maken: een beroep op onwenselijke gevolgen kan een deugdelijk argument zijn (maar dit hoeft niet) als het standpunt van het prescriptieve type is, terwijl dezelfde soort argumentatie altijd onredelijk is wanneer deze naar voren wordt gebracht voor een descriptief standpunt.²

Over de vraag of *ad populum*-argumentatie in bepaalde gevallen deugdelijk kan zijn, geeft de argumentatietheoretische literatuur geen eenduidig beeld, nog afgezien van het feit dat *ad populum*-argumentatie sowieso op vrij weinig diepgaande belangstelling kan rekenen. De analyses die er zijn, zijn vaak gericht op een bepaalde casus en bieden daardoor slechts *ad hoc* criteria voor de beoordeling. Zelfs in de uitgebreide monografie van Walton (1999) over dit argumentatietype ontbreekt een systematisch uitgewerkt en helder toepasbaar beoordelingsinstrumentarium. Het doel

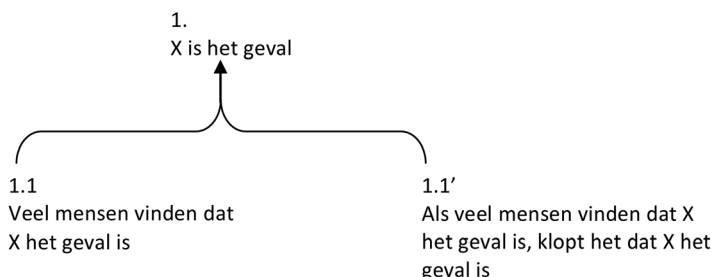
van deze bijdrage is dan ook om meer duidelijkheid te scheppen over de analyse en beoordeling van *ad populum*-argumentatie, waarbij ik deze argumentatie zal bestuderen in relatie tot elk van de drie in de pragma-dialectiek onderscheiden standpunten. Ik zal dit doen aan de hand van een kritische bespreking van de literatuur en een systematisering van de daarin verwoorde inzichten.

Een feitelijke en een prescriptieve variant van *ad populum*-argumentatie

Nolt (1984, p. 249) is een van de weinige auteurs die een onderscheid maakt tussen twee varianten van *ad populum*-argumentatie. Hoewel hij dit zelf niet expliciet zegt, hangt het onderscheid tussen beide typen samen met het type standpunt dat wordt ondersteund. Ten eerste onderscheidt Nolt een ‘feitelijke versie’ van *ad populum*-argumentatie, die gebaseerd is op een descriptief standpunt en die de volgende redenering behelst: ‘Believing that P is popular, so P’. In deze variant wordt uit het feit dat P een populair (dat wil zeggen door veel mensen gedragen) idee is, geconcludeerd dat P het geval is. Daarnaast beschrijft hij een ‘prescriptieve versie’, gebaseerd op een prescriptief standpunt: ‘Doing X is popular, so X is permissible (or should be done)’. In deze variant wordt uit het feit dat X een populaire bezigheid is, geconcludeerd dat X toelaatbaar is of uitgevoerd moet worden. Beide argumentaties zijn gebaseerd op de gedachte dat als veel mensen hetzelfde vinden of doen, het onwaarschijnlijk is dat zij het bij het verkeerde eind hebben. Volgens Nolt klopt deze gedachte niet en zijn beide varianten daarom drogredenen. Overigens lijkt Nolt de mogelijkheid open te houden dat *ad populum*-argumentatie soms redelijk kan zijn, aangezien hij beide varianten ‘generally fallacious’ (p. 249) noemt.³ Andere theoretici die net als Nolt binnen het raamwerk van de Informele Logica werken, beschouwen beide varianten – die zij overigens niet expliciet onderscheiden, maar die wel in de door hen besproken voorbeelden ontdekt kunnen worden – altijd als ondeugdelijk (Freeman 1995; Johnson & Blair 2006; Govier, 2010).

De door Nolt (1984) (en ook door anderen) genoemde gedachte die aan *ad populum*-argumentatie ten grondslag ligt – als veel mensen hetzelfde vinden of doen, is het onwaarschijnlijk dat zij het bij het verkeerde eind hebben – bepaalt de inferentie van deze argumentatie. Voor de feitelijke variant herformuleer ik deze gedachte tot de inferentie dat als veel mensen iets vinden, dit een aanwijzing is dat het klopt. Op basis hiervan kom ik tot de volgende reconstructie van de feitelijke variant, met in (1) het standpunt,

in (1.1) het argument waarin naar de mening van een grote groep mensen wordt verwezen en in (1.1') – het verzwegen argument – de geëxpliciteerde inferentie:



Figuur 1

Dat *ad populum*-argumentatie met een descriptief standpunt onredelijk is, wordt blootgelegd in het verzwegen argument.

Ad populum-argumentatie met een prescriptief standpunt zit ingewikkelder in elkaar dan *ad populum*-argumentatie met een descriptief standpunt. Dat komt door Nolts (1984) atypische omschrijving van deze variant, waarin niet naar de *mening* van een grote groep mensen wordt verwezen maar naar wat zij *doen*. Deze omschrijving lijkt ingegeven door het voorbeeld dat hij van deze variant geeft, namelijk een reclametekst van het type 'largest selling model in its class' (p. 250), die consumenten ertoe moet overhalen om het desbetreffende model te kopen. Dergelijke argumentatie heeft volgens Nolt twee implicaties. De eerste is dat als iedereen product X koopt, jij het ook moet doen. De tweede implicatie is dat omdat zoveel mensen het product kopen, het wel een goed product zal zijn. Dat de tweede implicatie een belangrijk onderdeel uitmaakt van *ad populum*-argumentatie in reclame blijkt uit het soort kritiek dat volgens Nolt op dergelijke argumentatie van toepassing is: 'The popularity of a product is often less an indicator of its quality than of the effectiveness of the advertising which promotes it'.

Dat *ad populum*-argumentatie in een reclame-context twee implicaties heeft, betekent dat deze argumentatie complexer in elkaar zit dan Nolts (1984) redeneerschema aangeeft. Dit type *ad populum*-argumentatie bestaat uit drie stappen: 1. Jij moet product X kopen, 2. want product X is goed, 3. want dat blijkt uit het feit dat veel mensen product X kopen. Als de redenering echter zo in elkaar zit, dan bevindt de *ad populum* – het beroep op wat veel mensen doen – zich op een subniveau in de redenering. Het (sub)standpunt dat met de *ad populum* wordt ondersteund is dan niet prescriptief maar evaluatief van aard ('X is goed'). De problematiek van dit

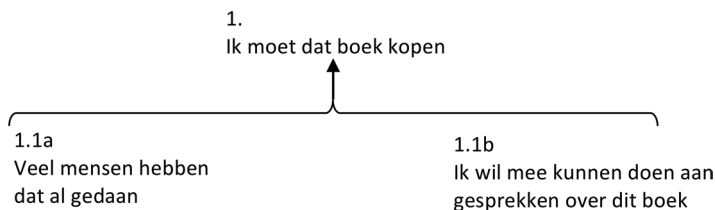
type standpunt alsmede de reconstructie van de bijbehorende *ad populum*-argumentatie zal ik in de volgende paragraaf behandelen.

Overigens vinden niet alle auteurs dat de bovengenoemde varianten van *ad populum*-argumentatie altijd onredelijk zijn. Walton (1999, pp. 102-102; 201-205; 2006, pp. 91-93) bespreekt een aantal voorbeelden waarin beide varianten wel redelijk zouden zijn. Zo kan de feitelijke variant van *ad populum*-argumentatie bijvoorbeeld gecombineerd worden met een autoriteitsargument, bijvoorbeeld wanneer de groep mensen naar wiens standpunt wordt verwezen een groep wetenschappers is. Hieronder volgt mijn eigen voorbeeld van zo'n gecombineerd *ad populum*-argument. In dit voorbeeld anticipeert argument (1.1b) op mogelijke kritiek die de inferentierelatie tussen argument (1.1a) en het standpunt ter discussie stelt:



Figuur 2

Ook de prescriptieve variant kan worden gecombineerd met een ander argument met als resultaat complexe *ad populum*-argumentatie die wel redelijk zou zijn (zie voor dit voorbeeld Walton, 1999, p. 34 e.v. en Kahane, 1984, p. 57):⁴



Figuur 3

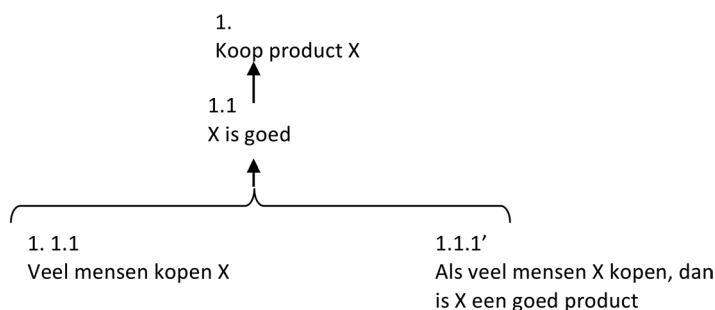
Zoals echter door Godden (2008) is opgemerkt, en overigens door Walton (1999) zelf ook wordt onderkend (pp. 205, 224), is er in deze gevallen geen sprake meer van de 'basisvorm' van *ad populum*-argumentatie. Het is immers juist de combinatie met dit extra argument die de argumentatie als geheel redelijk maakt.

De conclusie die tot dusver kan worden getrokken, is dat *ad populum*-argumentatie onredelijk is wanneer deze wordt aangevoerd voor een descriptief standpunt, tenzij er sprake is van een combinatie met een extra, nevenschikkend argument dat het onredelijke karakter aan de *ad populum* wegneemt. Over *ad populum*-argumentatie ter ondersteuning van een prescriptief standpunt is tot nu toe duidelijk geworden dat ook hier de combinatie met een extra argument redelijke argumentatie kan opleveren. Verder is gebleken dat als het prescriptieve standpunt zich voordoet in een reclame-context, er sprake is van complexe argumentatie met een evaluatief substandpunt; deze situatie zal nader worden geanalyseerd in de derde paragraaf. Verder zijn nog niet de auteurs aan bod geweest die aandacht besteden aan nog een ander type *ad populum*-argumentatie met een prescriptief standpunt, namelijk het inzetten van dergelijke argumentatie door politici ter ondersteuning van een beleidsvoorstel; ik zal dit type behandelen in de vierde paragraaf.

***Ad populum*-argumentatie ter ondersteuning van een evaluerend standpunt**

In de literatuur wordt nergens gesproken over *ad populum*-argumentatie ter ondersteuning van een evaluerend standpunt. Zoals echter in de vorige paragraaf bij de bespreking van Nolts (1984) prescriptieve variant is gebleken, zit er in prescriptieve *ad populum*-argumentatie in een reclamecontext een impliciet evaluerend substandpunt, namelijk ‘Product X is goed’. Nolt zelf lijkt de propositie ‘Product X is goed’ echter als descriptief in plaats van evaluerend te beschouwen. Dat blijkt uit het feit dat de tweede implicatie die hij toeschrijft aan *ad populum*-argumentatie in een reclame-context (zoals in de vorige paragraaf is beschreven), aanduidt als de feitelijke variant van deze argumentatie: ‘That’s a factual version’. In Nolts eigen woorden luidt deze implicatie: ‘since so many people buy the product, they must think it is good, hence it is good’ (p. 250). Omdat hij de eerste implicatie – ‘everyone else is buying this product, [so] you should too’ – prescriptief noemt, verbindt Nolt de labels *factual* en *prescriptive* kennelijk aan het consequens/standpunt van beide implicaties: ‘Jij moet product X kopen’ (prescriptief) versus ‘Product X is goed’ (feitelijk/descriptief). De uitspraak dat een product goed is, moet echter als evaluatief worden geduid, omdat deze geen objectief controleerbare stand van zaken weergeeft maar een subjectief waardeoordeel.

Ad populum-argumentatie in een reclamecontext bestaat dus uit complexe argumentatie waarin de verwijzing naar de grote groep mensen die een bepaald product koopt, zich bevindt op het subniveau van de argumentatie. Het prescriptieve standpunt op hoofdniveau ('Koop X') wordt ondersteund door een evaluerend substandpunt ('X is goed'), dat op zijn beurt wordt ondersteund door de verwijzing naar de vele mensen die X kopen. De inferentie die hoort bij het beroep op de hoeveelheid mensen is dat het feit dat veel mensen X kopen, een teken is dat X een goed product is:

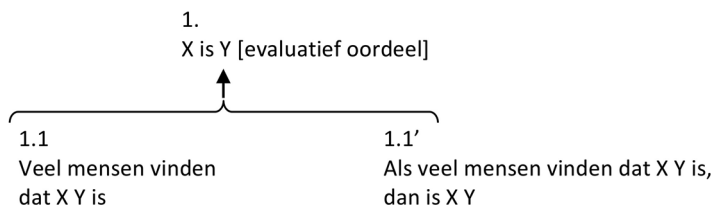


Figuur 4

Nu de argumentatie is gereconstrueerd, is het mijns inziens gerechtvaardigd de vraag te stellen of hier eigenlijk wel sprake is van *ad populum*-argumentatie en deze niet gewoon als reguliere kentekenargumentatie moet worden beschouwd. Het gaat in deze argumentatie immers om wat mensen doen, en niet om hun mening ter rechtvaardiging van een positief standpunt over die mening. Freeman, die ook reclame als voorbeeld geeft van (prescriptieve) *ad populum*-argumentatie, hanteert een definitie waarin wél wordt verwezen naar wat mensen vinden in plaats van wat mensen doen: '(...) an argument which claims that because many, most, all people (...) approve a certain course of action, we have sufficient reason to accept that (...) course of action as right' (1995, pp. 267-268). Dus als veel mensen een bepaalde handeling goedkeuren, is het een goede handeling. Toegepast op een reclamecontext is deze definitie echter ook problematisch, want welke *course of action* moet worden goedgekeurd – is dat het kopen van product X? Dan zou volgens deze definitie de *ad populum* in een reclameadvertentie de volgende, vreemde redenering opleveren: Koop product X, want X kopen is goed, want veel mensen keuren het goed om X te kopen.

De vraag of een beroep op wat mensen doen (in plaats van vinden) wel *ad populum*-argumentatie betreft, is een definitiekwestie die ik verder terzijde laat. In de onderstaande reconstructie van *ad populum*-argumentatie

ter ondersteuning van een evaluerend standpunt ga ik uit van een definitie van deze argumentatie waarin een beroep wordt gedaan op wat een grote groep mensen *vindt*:



Figuur 5

Hoe zit het met de deugdelijkheid van *ad populum*-argumentatie ter ondersteuning van een evaluerend standpunt? Klopt de aanname dat – zoals uitgedrukt in argument (1.1') – als een grote groep mensen een bepaald waardeoordeel heeft, dit oordeel correct is? Wie bepaalt eigenlijk wat goed, slecht, mooi, lelijk enzovoort is? In *NRC Handelsblad* stond ooit een ingezonden brief waarin werd geklaagd over de elitaire smaak van filmcritici, die in hun top-vijf-lijstjes geen enkele film hebben staan die de briefschrijver en film liefhebbers uit zijn omgeving kon behagen, en vice versa. De brief sluit af met:

Film hoeft niet altijd kunst te zijn, maar het is wel een hele kunst om een goede film te maken. Ik laat me liever bijna drie uur vermaken door de nieuwe Harry Potter in plaats van half slapend naar een zwijgende film uit Zuid-Korea te kijken. Dat is persoonlijk, maar de miljoenen die wereldwijd ook naar Potter gaan kunnen toch niet allemaal een foute smaak hebben? (*NRC Handelsblad*, 3 januari 2006)

Het probleem met evaluerende standpunten is dat het überhaupt lastig is om deze te rechtvaardigen. In tegenstelling tot descriptieve proposities kunnen proposities die een ethisch of esthetisch oordeel uitdrukken niet 'bewezen' worden; er bestaan geen objectieve criteria voor. Dat neemt natuurlijk niet weg dat er wel degelijk op een rationele manier over dergelijke proposities kan worden gediscussieerd, onder andere door als voorwaarde te stellen dat de argumentatie gebaseerd moet zijn op gemeenschappelijke uitgangspunten. Dit betekent dat in discussies over ethische of esthetische kwesties, afhankelijk van wie er overtuigd moet worden, een beroep zal worden gedaan op waarden en normen die berusten op (geëvolueerde) consensus binnen de desbetreffende maatschappij of subcultuur.⁵

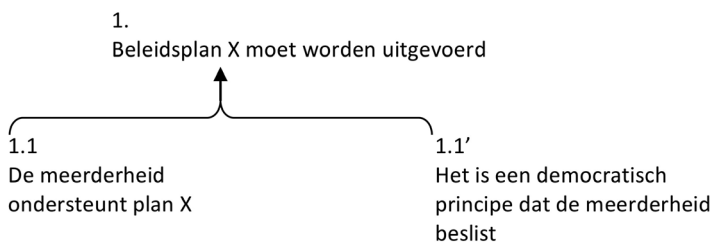
Als argumentatie voor een waardeoordeel echter gegrond is in consensus, dan is het een begrijpelijke vraag waarom de uitgangspunten van een select gezelschap filmcritici betere argumentatie zou opleveren voor een evaluerend standpunt over een film dan de consensus van ‘miljoenen’ mensen die zich met de desbetreffende film hebben vermaakt. Als esthetische criteria immers consensus vereisen, waarom volstaat dan niet de brede consensus die met een *ad populum*-argument wordt uitgedrukt? Het probleem met de in *ad populum*-argumentatie uitgedrukte consensus is echter dat in dergelijke argumentatie de overwegingen van de groep mensen waarnaar wordt verwezen altijd ongearticuleerd blijven. Ik bedoel daarmee niet dat deze groep mensen (bijvoorbeeld de Harry Potter-liefhebbers) geen criteria zou hebben of toepassen; ik bedoel daarmee dat het inherent is aan *ad populum*-argumentatie dat daarin geen beroep wordt gedaan op die criteria, omdat deze argumentatie – bij definitie – slechts bestaat uit een verwijzing naar de mening van een groep mensen en niet naar de argumenten die ze voor die mening hebben. Een beroep op wat mensen vinden biedt dus geen materiële ondersteuning voor een standpunt, en dat is de belangrijkste reden voor de onredelijkheid van deze argumentatie.

Er is ook nog een ander probleem met de in *ad populum*-argumentatie uitgedrukte consensus. Consensus is namelijk alleen van waarde voor de leden van de groep waarbinnen die consensus bestaat. Voor leden van een groep met andere uitgangspunten heeft die consensus helemaal geen betekenis. De briefschrijver lijkt te wensen dat objectief vastgesteld kan worden dat Harry Potter-liefhebbers geen foute smaak hebben. Die objectieve maatstaf is er echter niet. Alleen binnen een bepaald waardensysteem bestaan er criteria die binnen dat systeem wellicht objectief genoemd kunnen worden. Maar zodra je je in een ander waardensysteem bevindt, kunnen er weer heel andere ‘objectieve’ criteria gehanteerd worden.⁶ Er is geen objectief criterium om het ene waardensysteem boven het andere te verkiezen, ook niet de macht van het getal.

Procedurele *ad populum*-argumentatie

In een laatste variant van *ad populum*-argumentatie waaraan in de literatuur aandacht wordt besteed, ondersteunt deze een standpunt waarin politiek beleid wordt aan- of afgeraden. Een aantal auteurs pleit ervoor om dergelijke argumentatie als in potentie redelijk te beschouwen. Volgens Minot (1981, p. 230) is publieke opinie immers de basis voor besluitvorming: ‘The desire of the populace is the means for deciding an issue’. Ook Andone

(2015), die een geval van *ad populum*-argumentatie van een Europese politica analyseert, en Oswald & Hart (2013), die onder andere de cognitieve effecten van *ad populum*-argumentatie beschrijven, zijn van mening dat dergelijke argumentatie deugdelijk kan zijn. In beide publicaties vormt het feit dat een grote groep mensen bepaald beleid voorstaat, een democratisch gemotiveerde reden om dat beleid in te voeren.⁷ De inferentie van dit type *ad populum*-argumentatie bestaat dan ook uit de verwoording van deze democratische beslisregel en dus niet uit de gedachte dat als veel mensen iets vinden het wel zal kloppen:



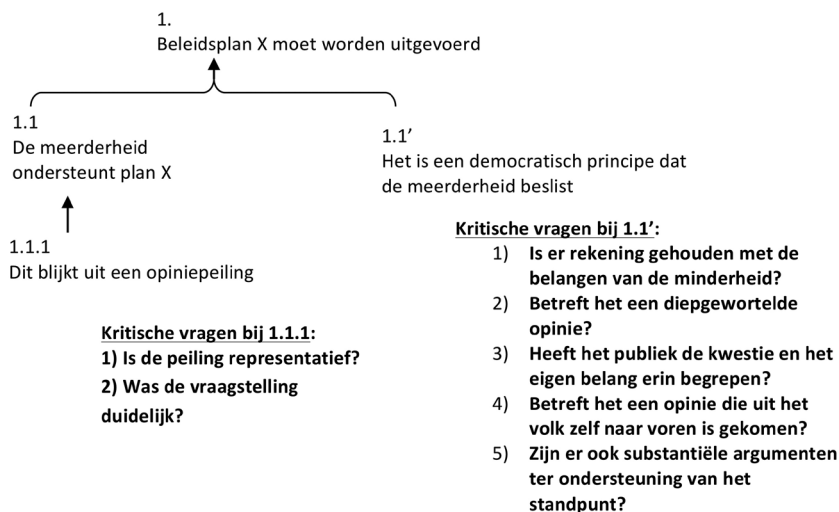
Figuur 6

Maakt het beroep op een democratische beslisregel *ad populum*-argumentatie inderdaad redelijk? Toen toenmalig minister Donner in 2006 stelde dat de sharia zou moeten worden ingevoerd mocht tweederde van de Nederlandse bevolking hier voorstander van zijn, viel iedereen over hem heen.⁸ Want hoewel we geneigd zijn om een groot gewicht toe te kennen aan meerderheidsopvattingen wanneer deze stroken met onze eigen opvattingen, vinden we de mening van de meerderheid problematisch als we zelf een andere mening hebben. Overigens geeft dit voorbeeld aanleiding om een onderscheid te maken tussen politieke besluitvorming als gevolg van een meerderheid in het parlement of bij een (bindende) referendumuitslag enerzijds en anderzijds het beroep op wat mensen vinden als argument in een discussie. In het eerste geval is het natuurlijk wel degelijk zo dat de uitslag van de stemming de richting van het politieke beleid bepaalt. Donner meende dus terecht dat als het democratische besluitvormingsproces resulteert in een meerderheidssteun voor invoering van sharia, die invoering een feit zal zijn. Over besluitvorming die het gevolg is van het tellen van stemmen gaat dit artikel echter niet. Wat hier centraal staat, is het beroep op wat een grote groep mensen vindt ter rechtvaardiging van een standpunt in een discussie over een beleidskwestie, en de vraag of zo'n argument deugdelijk kan zijn. Dus stel dat er een maatschappelijke discussie zou zijn over mogelijke invoering van de sharia, dan zou het een argument kunnen

zijn dat dit plan moet worden uitgevoerd omdat veel mensen voor zijn. Een dergelijk argument zou gebaseerd kunnen zijn op een vermoeden van wat mensen van een bepaalde kwestie vinden of op een peiling daarvan.

Hoe redelijk is dan *ad populum*-argumentatie ter rechtvaardiging van een standpunt in een discussie over een beleidskwestie – een variant van *ad populum*-argumentatie die ik vanwege de inferentie verder ‘procedurele *ad populum*-argumentatie’ zal noemen? In de literatuur wordt een aantal overwegingen genoemd die geïnterpreteerd kunnen worden als criteria voor de beoordeling van dit type *ad populum*-argumentatie. Oswald & Hart (2013) leiden uit hun bespreking van een bepaalde casus af dat een voorwaarde voor redelijke *ad populum*-argumentatie in een politieke context in ieder geval is dat aangetoond kan worden dat de genoemde groep mensen de genoemde opinie inderdaad heeft, bijvoorbeeld met opiniepeilingen. Ook Walton (1999) bespreekt her en der een aantal zaken die als beoordelingscriteria kunnen worden gebruikt; in zijn eigen beoordelingsinstrumentarium voor *ad populum*-argumentatie zijn ze echter niet opgenomen.⁹ In de eerste plaats zegt hij dat als in *ad populum*-argumentatie een beroep wordt gedaan op opiniepeilingen, de hiervoor uitgevoerde steekproef representatief moet zijn en geen dubbelzinnige of vage vraagstelling mag bevatten (pp. 2-10, p. 125). Verder merkt hij op dat meerderheidsbesluitvorming niet mag ontaarden in de tirannie van de meerderheid (p. 11; zie al eerder Toulmin, Rieke, & Janik, 1984, p. 146). Essentieel voor een democratie is immers ook dat rekening wordt gehouden met minderheidsbelangen. Een laatste criterium dat aan Walton kan worden ontleend is dat publieke opinie volatiel is: deze kan gemakkelijk omslaan, en is daarom een minder goede motivatie voor de ontwikkeling van beleid (p. 11; zie ook Johnson & Blair, 2006, p. 179).

Bij andere theoretici kunnen nog enkele andere criteria worden gevonden. Zo zegt Minot (1984, pp. 230-231) dat het centrale beoordelingspunt bij *ad populum*-argumentatie ter ondersteuning van een beleidsstandpunt is of het publiek de desbetreffende beleidskwestie waarover het een opinie heeft wel goed heeft begrepen. Müller (2006, p. 10), onderzoeker van populisme, merkt op dat het vaak politici zelf zijn die de publieke opinie eerst beïnvloeden en de beïnvloede opinie later aanvoeren als argumentatie voor de door hen voorgestelde maatregelen. Als voorbeeld haalt hij een uitspraak van Erdogan aan uit juli 2016, na de coup, waarin hij de doodstraf verdedigt door te zeggen dat ‘zijn’ volk dat wil, terwijl hij het plan eerst zelf had geïntroduceerd en verdedigd. Al met al kunnen de inzichten uit de literatuur worden samengebracht in het volgende analyse- en beoordelingsinstrumentarium voor procedurele *ad populum*-argumentatie:



Figuur 7

Omdat het beroep op de mening van de meerderheid idealiter wordt bear-gumenteed met cijfers – zoals gesteld door Oswald & Hart (2013) – is deze overweging als argument (1.1.1) opgenomen. Bij dit argument zijn de door Walton genoemde criteria ter beoordeling van peilingen opgenomen als kritische vragen. De kritische vragen bij verzwegen argument (1.1') stellen het democratische principe dat de meerderheid beslist ter discussie. Alle vragen reflecteren de in de literatuur gevonden overwegingen. De onderste van deze vragen is hierboven nog niet besproken. Deze representeert het idee dat een procedureel *ad populum*-argument weliswaar niet inherent onredelijk is, maar – gezien de in de literatuur gemaakte kanttekeningen bij dit argument – anderzijds ook niet volstaat als enige argument in een discussie over beleid. Dit brengt me ertoe te concluderen dat een beleids-standpunt altijd mede ondersteund zal moeten worden met inhoudelijke argumenten.

Conclusie

In deze bijdrage heb ik de literatuur over *ad populum*-argumentatie kritisch besproken en de daarin naar voren komende inzichten gesystematiseerd en aangevuld. Als resultaat heb ik reconstructies gepresenteerd van diverse va-rianten van *ad populum*-argumentatie, die inzicht bieden in de redelijkheid of de onredelijkheid van de desbetreffende variant. Van de in de literatuur

onderscheiden descriptieve en prescriptieve variant is de descriptieve variant – waarmee een descriptief standpunt wordt ondersteund – alleen deugdelijk in combinatie met aanvullende nevenschikkende argumentatie. De prescriptieve variant – waarmee een prescriptief standpunt wordt ondersteund – wordt in de literatuur zowel aan een reclamecontext als aan een politieke besluitvormingscontext gekoppeld. *Ad populum*-argumentatie in een reclamecontext heb ik geanalyseerd als complexe argumentatie waarin het beroep op een grote groep mensen een evaluatief (sub)standpunt ondersteunt. Ik heb betoogd dat hoewel er voor evaluatieve standpunten geen objectieve criteria bestaan en de overtuigingskracht van argumentatie voor een waardeoordeel juist lijkt af te hangen van haar basis in uitgangspunten die door een grote groep mensen worden gedeeld, *ad populum*-argumentatie ter ondersteuning van zo'n standpunt toch inherent onredelijk is. *Ad populum*-argumentatie in de context van politieke besluitvorming – door mij 'procedurele *ad populum*-argumentatie' genoemd – is niet inherent onredelijk, maar moet desondanks worden gekarakteriseerd als een argument dat te zwak is om zelfstandig de bewijslast voor het standpunt te dragen. De hypothese dat het standpunt van invloed is op de beoordeling van *ad populum*-argumentatie is daarmee in zoverre bevestigd dat alleen een prescriptief standpunt geen inherente onredelijkheid oplevert, mits dit standpunt een beleidskwestie betreft waaraan een procedurele beslisregel ten grondslag ligt.

Noten

1. Ik ben de twee anonieme reviewers van dit artikel erkentelijk voor hun waardevolle commentaar.
2. In het eerste geval is er sprake van pragmatische argumentatie en in het tweede geval van de *ad consequentiam*-drogreden.
3. Op p. 250 schrijft hij dat als de *ad populum*-argumentatie niet berust op de aanname dat de groep mensen waarnaar wordt verwezen betrouwbaar ('reliable') is, dit type argumentatie sowieso drogredelijk is. Als deze aanname wel van kracht is, is de argumentatie 'seldom strong'.
4. Bij dit voorbeeld moet worden opgemerkt dat het feit dat veel mensen een boek hebben gekocht, niet wil zeggen dat al deze mensen het ook gelezen hebben, waardoor de argumentatie toch nog aanvechtbaar is. Waltons punt is evenwel dat de argumentatie in deze constellatie niet meer als inherent onredelijk kan worden beschouwd.
5. Ook de esthetische criteria van de kunstwereld, die wellicht als meer objectief worden beschouwd omdat ze zijn ontwikkeld door kunstexperts, bestaan bij de gratie van consensus binnen die wereld.

6. Overigens kan dit probleem ook spelen in discussies over descriptieve proposities: ook dan kan het zo zijn dat binnen verschillende (sub)culturen verschillende maatstaven worden aangelegd. Een voorbeeld hiervan is de frictie tussen religie en wetenschap.
7. Ook Johnson en Blair (2006, p. 179) merken op dat het beleid dat in een land wordt ontwikkeld niet te veel mag afwijken van wat de mensen in een land vinden. Zij verbinden hier echter niet de conclusie aan dat dit redelijke *ad populum*-argumentatie kan opleveren. *Ad populum*-argumentatie is volgens hen namelijk altijd gebaseerd op een 'criteriële' regel, wat wil zeggen dat de mening van de meerderheid het criterium vormt voor de waarheid of de aannemelijkheid van die mening (p. 178). Zij erkennen wel dat in een politieke context een procedurele beslisregel een rol kan spelen, maar zien deze alleen toegepast wanneer gevolg wordt gegeven aan de uitslag van bijvoorbeeld een referendum.
8. Dit zei hij daarover in zijn boek *Het land van haat en tijd*: 'De meerderheid telt. Dat is nou juist de essentie van democratie.' Kamervragen brachten hem ertoe te zeggen dat het om een puur theoretische stelling ging en dat hij er geen misverstand over wilde laten bestaan dat hij een sharia in Nederland afkeurt.
9. Waltons 'officiële' (want expliciet zo gepresenteerde) beoordelingsinstrument is vrij algemeen van aard. In hoofdstuk 8 presenteert hij 'The four steps of an evaluation' van een *ad populum*: 1) beoordeel de aanvaardbaarheid van de premissen, 2) beoordeel de dialectische relevantie van het argument aan de hand van het type dialoog, 3) beoordeel de sterkte van het argument in relatie tot andere argumentatie die ermee samenhangt (hiermee wordt aan de in de tweede paragraaf van dit artikel nevenschikkende argumentatie gerefereerd), 4) beoordeel hoe de argumentatie appelleert aan de gebondenheden van het publiek (bijvoorbeeld: wordt er ruimte gelaten om kritische vragen te stellen?) (pp. 250-252). Wat in elke stap concreet moet worden nagegaan en wanneer toepassing ervan tot een positief of een negatief oordeel leidt, wordt helaas niet duidelijk. In hoofdstuk 9 gaat Walton nog wel specifiek in op een aantal subtypen van *ad populum*-argumentatie, waaronder het type 'public opinion polls'. Hij benadrukt hier het belang van de evaluatiestappen 1 en 2, maar toepassing op het desbetreffende subtype blijft achterwege. In hoofdstuk 7 wordt wel concreet iets gezegd over beoordelingsstap 2, namelijk dat *ad populum*-argumentatie alleen irrelevant is wanneer deze gebruikt wordt in een 'inquiry type of dialogue' (pp. 188-194).

Bibliografie

- Andone, C. (2015). Engagement et non-engagement dans les appels à la majorité par des hommes politiques. *Argumentation et analyse du discours*, 15 (13 pp.).
- Eemeren F.H. van (2010). *Strategic Maneuvering in Argumentative Discourse. Extending the Pragma-dialectical Theory of Argumentation*. Amsterdam/Philadelphia, PA: John Benjamins.
- Freeman, J.B. (1995). The appeal to popularity and presumption by common knowledge. In H.V. Hansen & R.C. Pinto (Eds.), *Fallacies: Classical and Contemporary Readings*. University Park, PA: Pennsylvania State University Press.
- Godden, D. (2008). On common ground and *ad populum*: Acceptance as grounds for acceptability. *Philosophy and Rhetoric*, 41, 101-129.
- Govier, T. (2010). *A Practical Study of Argument* (7th ed.). Belmont, CA: Wadsworth.
- Johnson, R.H., & Blair, J.A. (2006). *Logical Self-Defense*. New York, NY: Idea Press.

- Kahane, H. (1980), *Logic and Contemporary Rhetoric: The Uses of Logic in Everyday Life* (3rd ed.). Belmont, CA: Wadsworth.
- Minot, W.S. (1981). A rhetorical view of fallacies: Ad hominem and *ad populum*. *Rhetoric Society Quarterly*, 11, 222-235.
- Müller, J.W. (2016). *Wat is populisme?* Amsterdam: Nieuw Amsterdam.
- Nolt, J.E. (1984). *Informal Logic. Possible Worlds and Imagination*. New York, NY: McGraw-Hill Book Company.
- Oswald, S., & Hart, C. (2013). Trust based on bias: Cognitive constraints on source-related fallacies. In D. Mohammed & M. Lewinski (Eds.), *Virtues of Argumentation. Proceedings of the 10th International Conference of the Ontario Society for the Study of Argumentation (OSSA), 22-26 May 2013*. Windsor, ON: OSSA (pp. 1-13).
- Toulmin, S., Rieke, R.D., & Janik, A. (1984). *An Introduction to Reasoning*. New York, NY: MacMillan.
- Walton, D.N. (1999). *Appeal to Popular Ppinion*. University Park, PA: Pennsylvania State University Press.
- Walton, D.N. (2006). *Fundamentals of Critical Argumentation*. Cambridge: Cambridge University Press.

Pressie en argumentatie in maatschappelijke discussies

Jan Albert van Laar en Erik C. W. Krabbe
Rijksuniversiteit Groningen

TVT 41 (1): 167–179

DOI: 10.5117/TVT2019.1.011.ALBE

Abstract

Pressure and argumentation in public controversies

Discussants in a public controversy are also players in a social arena. Under what circumstances can means of exerting pressure promote reasonable outcomes, and when are they rather hindrances to their achievement? We discuss the use of persuasion dialogues and negotiation dialogues by those who aim at a reasonable outcome. We show how in a public controversy both types of dialogue can be intertwined. Then, we examine in what ways participants in a public controversy can exert pressure on others. Finally, we discuss whether these means of pressure hinder the achievement of a reasonable outcome, both from the viewpoint of the pursuit of a resolution as well as from that of the pursuit of a compromise. We conclude that exerting pressure need not degenerate into committing a fallacy of either the type *argumentum ad baculum* or the type *argumentum ad carotam*.

Keywords: argumentation, fallacies, negotiation dialogue, persuasion dialogue, pressure, public controversy

Inleiding

Partijen in een maatschappelijke discussie zijn ook spelers in een maatschappelijk krachtenveld. De deelnemers argumenteren om anderen van hun gelijk te overtuigen maar oefenen ook druk uit om hun zin te krijgen. Wij richten ons op het uitoefenen van pressie door af te schrikken met

sancties of juist te verlokken met beloningen. Wanneer kunnen dergelijke pressiemiddelen een redelijke uitkomst van een maatschappelijke discussie bevorderen en wanneer staan ze die alleen maar in de weg?

We proberen deze vraag te beantwoorden vanuit een dialogisch gezichtspunt: een maatschappelijke discussie zien we als een samenstel van onder meer overtuigingsdialogen en onderhandelingsdialogen.¹ In een overtuigingsdialoog zijn de deelnemers gericht op een inhoudelijke oplossing. In een onderhandelingsdialoog zijn de deelnemers gericht op het vinden van een voor ieder acceptabel compromis. In beide gevallen kan er sprake zijn van een redelijke uitkomst.

In het gunstigste geval speelt in een overtuigingsdialoog pressie geen rol. Maar pressie is inherent aan onderhandelen. In de praktijk lopen overtuigingsdialogen en onderhandelingsdialogen door elkaar zodat de pressie in een onderhandelingsdialoog ook invloed kan hebben op een overtuigingsdialoog. Met het oog hierop willen we bijdragen aan de ontwikkeling van instrumenten voor de ontleding en beoordeling van argumentatie in maatschappelijke krachtmetingen, en richten ons daarbij op de rol van voorwaardelijke sancties en voorwaardelijke beloningen (Amgoud & Prade, 2006; Van Laar & Krabbe, 2016a, 2016b).

In de volgende paragraaf gaan we nader in op de overtuigingsdialogen en onderhandelingsdialogen voor hen die streven naar een redelijke uitkomst van een maatschappelijke discussie. Daarna gaan we na op welke manier je in een maatschappelijke discussie pressie kan uitoefenen. In de daaropvolgende paragraaf bespreken we of en hoe deze pressiemiddelen een redelijke uitkomst hinderen of bevorderen. We concluderen dat pressie wel kan, maar niet hoeft te ontaarden in het begaan van een drogreden, zoals een *argumentum ad baculum* ('met de stok') of juist een *argumentum ad carotam* ('met de wortel').

We illustreren onze bevindingen met voorbeelden uit de in Nederland gevoerde maatschappelijke discussie over de energietransitie.²

Overtuigingsdialogen en onderhandelingsdialogen in een maatschappelijke discussie

Een maatschappelijke discussie vormt een mozaïek van diverse soorten dialogen, en het uitgangspunt ervan is doorgaans een combinatie van: (a) conflicterende belangen (b) meningsverschillen en (c) problemen waarbij nog niemand een standpunt heeft ingenomen. De centrale rol van het uitwisselen van overwegingen duidt op een gemeenschappelijk streven naar

een gezamenlijk onderschreven uitkomst die een kritische toetsing kan weerstaan, en die we aanduiden als een *redelijke uitkomst*.

We richten ons op de overtuigingsdialoog en de onderhandelingsdialoog. In een overtuigingsdialoog proberen deelnemers een verschil van mening op basis van de merites van de kwestie *op te lossen* door argumentatie en argumentatieve kritiek uit te wisselen. Maar deelnemers aan een maatschappelijke discussie zijn vaak ook spelers in het maatschappelijke krachtenveld die proberen in een onderhandelingsdialoog een belangenconflict te *beslechten* met een redelijk compromis dat de weerslag vormt van een uitwisseling van bod en tegenbod (Van Eemeren & Grootendorst, 2000, p. 62). Een redelijke uitkomst, van een maatschappelijke discussie of van een deel ervan, kan bestaan uit een *oplossing*, maar ook uit een *compromis*, waarbij – terwijl de meningen blijven verschillen – de partijen bereid zijn genoeg te nemen met een beleidsvoorstel dat in hun ogen niet in alle opzichten voldoet maar dat toch te verkiezen valt boven een situatie waarin er geen compromis is bereikt. Ook in dergelijke gevallen heeft de uitkomst een soort van kritische toetsing ondergaan.

Soms ligt het voor de hand om een bijdrage te beoordelen vanuit normen voor overtuigingsdialogen. Soms vanuit normen voor onderhandelingsdialogen. Maar bij fragmenten waarin een overtuigingsdialoog en een onderhandelingsdialoog met elkaar zijn verweven is het zowel gerechtvaardigd om de inbreng te beoordelen vanuit een overtuigingsperspectief als vanuit een onderhandelingsperspectief, al naar gelang de interesse van de beoordelaar. Beide soorten dialoog vragen een ander gezichtspunt als het gaat om de beoordeling van pressiemiddelen.

In de regel gaat ook onderhandelen gepaard met argumenteren (Van Laar & Krabbe, 2016c, 2018). Neem een karakteristiek bod, zoals “Je kunt deze piano kopen voor slechts 6000 Euro!” Een dergelijk voorwaardelijk bod is van de vorm: “Als jij bereid bent mij concessies X te doen dan en dan alleen ben ik bereid jou concessies Y te doen.” Zo’n bod kan zo begrepen worden dat er ook een “profijtelijsheidsargumentatie” mee wordt uitgedrukt: “Het is profijtelijk voor jou om mijn bod te aanvaarden. Immers, deze piano is jou minstens 6000 Euro waard, en ik vraag slechts 6000 Euro.” Een tegenbod, zoals “Nee, maar ik ben wel bereid je er 5000 Euro voor te geven!” drukt een kritische reactie op de eerdere profijtelijsheidsargumentatie uit, maar bevat ook zelf een nieuwe overtuigingspoging. Profijtelijsheidsargumentatie is een toepassing van het volgende argumentatieschema: “Aanvaarding van mijn bod is profijtelijk voor jou, want als, en alleen als, je mijn bod aanvaardt verwerf je X ten koste van Y, en X ten koste van Y is profijtelijk voor

jou". Dit schema is een variant op het schema voor pragmatische argumentatie (Van Eemeren, 2016, p. 17).³

Een typische onderhandelingsdialoog bestaat uit een uitwisseling van bod en tegenbod. Maar met ieder bod kan zo'n profijtelijkheidsargumentatie geassocieerd worden. Een profijtelijkheidsargumentatie valt vervolgens te begrijpen als een bijdrage aan een overtuigingsdialoog waarin degene die het bod heeft gedaan eventuele twijfels bij de ontvanger ten aanzien van de profijtelijkheid van het betreffende bod met argumentatie probeert weg te nemen. Deze overtuigingsdialogen zijn daarmee *ingebed* in de onderhandelingsdialoog. Zoals we in de volgende paragraaf zullen zien is inbedding een vorm van verwevenheid van overtuigings- en onderhandelingsdialogen, maar niet de enige.

Pressie komt in de argumentatietheorie vooral aan de orde bij de bespreking van de drogreden *argumentum ad baculum* (Walton, 2000, 2014). Maar je kunt mensen ook onder druk zetten door ze te verlokken met een beloning. Dus willen we in navolging van Woods ook de drogreden *argumentum ad carotam* aan de orde stellen (Woods, 2004, p. 80). In beide gevallen gaat het in onze terminologie om bijdragen waarin iemand onder druk wordt gezet op een wijze die indruist tegen de normen die gelden voor de overtuigingsdialoog. Dat een bijdrage vanuit het overtuigingsperspectief illegitiem is, zegt trouwens nog niet veel over de legitimiteit van die bijdrage vanuit het onderhandelingsperspectief. Beide perspectieven hebben immers hun eigen normativiteit (Walton & Krabbe, 1995, Walton, 1998).

De term 'drogreden' reserveren we, zoals in de pragma-dialectiek, voor illegitieme zetten in een overtuigingsdialoog (Van Eemeren & Grootendorst, 2000, 2004). We gaan uit van de volgende dialectische definitie van de drogreden *argumentum ad baculum*: Discussiant A begaat met bijdrage *c* tegenover discussiant B de drogreden *argumentum ad baculum* dan en slechts dan als A met *c* aan B duidelijk maakt dat A ertoe over zal gaan B te straffen als B zijn of haar bijdrage aan de overtuigingsdialoog (standpunt, kritiek, argumentatie, kritische vraag, etc.) niet opgeeft. Daarmee volgen we grotendeels de pragma-dialectische drogredentheorie volgens welke deze drogreden een overtreding vormt van de Vrijheidsregel (Van Eemeren & Grootendorst, 2000, p. 191, 2004, p. 190). Discussiant A probeert immers met het dreigement in bijdrage *c* B de mond te snoeren. We merken hierbij op dat dergelijke overtredingen van de Vrijheidsregel o. i. niet alleen voorkomen in de confrontatiefase van de discussie waarin het initiële verschil van mening wordt gearticuleerd, maar ook tijdens de uitwisseling van argumentatie en kritiek.

We stellen voor de drogreden *argumentum ad carotam* op parallelle wijze te definiëren. Discussiant A begaat tegenover discussiant B met bijdrage *c* de drogreden *argumentum ad carotam* dan en slechts dan als A met *c* aan B duidelijk maakt dat A ertoe over zal gaan B te belonen als B zijn of haar bijdrage aan de overtuigingsdialoog (standpunt, kritiek, argumentatie, kritische vraag, etc) opgeeft. Ook het *argumentum ad carotam* overtreedt de vrijheidsregel en vormt een drogreden die tevens buiten de confrontatiefase kan worden begaan.

De rol van pressie in de maatschappelijke discussie over de energietransitie

Onderhandelingsdialogen en overtuigingsdialogen zijn in een maatschappelijke discussie op verschillende manieren verweven. Sommige bijdragen aan een onderhandelingsdialoog bevatten, zoals uitgelegd in de vorige paragraaf, profijtelijsheidsargumentatie die deel uit maakt van een overtuigingsdialoog welke in de onderhandelingsdialoog is ingebed. Maar niet alleen kan een dialoog van het ene type zijn ingebed in een dialoog van het andere type, ook zijn er hybride bijdragen: deze bevatten niet alleen een zet in een onderhandelingsdialoog maar ook een argumentatie in een overtuigingsdialoog. Verder zijn er passages die *dubbel gebruikt* worden, aangezien ze een functie vervullen in beide dialoogtypen.

Voorbeeld 1 illustreert de hybride bijdrage.

De bedrijven, waaronder Siemens, Eneco, Shell, Heineken, Schiphol, Van Oord en het Havenbedrijf Rotterdam, doen hun oproep woensdag in een brief in de Volkskrant. Zij ondersteunen daarmee het initiatief van Samsom en Klaver om een nationale Klimaatwet te maken. (...)

‘Wij zijn ervan overtuigd dat de energietransitie moet plaatsvinden om klimaatverandering tegen te gaan en zien het versnellen van de transitie ook als een kans voor het ontwikkelen van een nieuwe economie’, schrijven de bedrijven. (...)

Bedrijven hebben dan wel behoefte aan duidelijkheid ‘over de kabinetten heen’, betogen ze. Daarom moet het in een wet geregeld worden. ‘De afspraken die we maken moeten bindend zijn, want daar baseren we onze investeringen op.’ (Du Pré, 2016, p. 2).

Voorbeeld 1 Transitie is ook een kans

In deze hybride bijdrage argumenteren de bedrijven enerzijds voor de stelling dat de klimaatwet er moet komen, aangezien de energietransitie versneld moet worden. Anderzijds wekken ze ten minste de schijn de overheid een onderhandelingsbod te doen: in ruil voor een langdurige verplichting van de politiek middelen een klimaatwet zijn de bedrijven bereid meer te investeren.

Voorbeeld 2 illustreert de dubbel gebruikte passage.

Vieze energie van kolen is, alle kosten meegerekend, veel duurder dan schone energie. Helaas worden alle kosten nog niet meegenomen in de prijs. Daarom moeten we de markt soms een handje helpen. D66 wil alle kolencentrales in Nederland zo snel mogelijk, maar uiterlijk in 2025, op verantwoorde wijze sluiten – beginnend bij de meest vervuilende. We stellen de stroomvoorziening veilig door de groei van het aandeel van duurzame stroom, de inzet van bestaande gascentrales en door goede transportverbindingen met ons omliggende landen. (D66, 2016, p. 32)

Voorbeeld 2 D66 wil kolencentrales sluiten

In deze dubbel gebruikte passage argumenteert D66 voor het sluiten van de kolencentrales, en draagt daarmee bij aan een overtuigingsdialoog over deze kwestie. Maar aangezien dit een fragment is uit het verkiezingsprogramma is de achterliggende boodschap tegelijkertijd dat deze beleidsvoorkeur een eis zal zijn in eventueel te voeren coalitiebesprekingen.

Het soort pressie dat kenmerkend is voor een onderhandelingsdialoog kan van invloed zijn op een in dezelfde context gevoerde overtuigingsdialoog. Een partij kan gemotiveerd zijn om de overtuigingsdialoog in een bepaalde richting bij te sturen om te voorkomen dat onwelgevallige resultaten van de overtuigingsdialoog in de onderhandelingsdialoog haar keuzeruimte beperken of haar onderhandelingspositie verzwakken. Als bijvoorbeeld beide partijen het in een overtuigingsdialoog erover eens worden dat de klimaatdoelen niet gehaald kunnen worden zonder alle kolencentrales binnen vijf jaar te sluiten, dan kan in een onderhandelingsdialoog de conservatievere partij niet langer met goed fatsoen van de groenere partij de concessie vragen om in te stemmen met het pas sluiten van de laatste kolencentrale over tien jaar. Daarom kan een partij geneigd zijn om de pressie die in de onderhandelingsdialoog normaal is tevens te gebruiken om de overtuigingsdialoog bij te sturen. De conservatievere partij, bijvoorbeeld, kan zo in een overtuigingsdialoog druk uitoefenen op de groenere partij om te voorkomen dat ze moet toegeven dat een snelle sluiting van de kolencentrales nodig is.

In onderhandelingsdialogen is pressie normaal. Dat komt doordat je met een bod altijd een bepaald soort macht uitoefent: het verlenen van de concessies Y is een lokkertje om concessies X los te krijgen; het niet verlenen van concessies Y is een sanctie voor de weigering van concessies X.⁴ Zo gezien bevat ieder bod een “dreigement”: de beloofde concessies Y gaan niet door als de ander niet over de brug komt met X.

Met dergelijke pressie wijzig je de sociale situatie waarbinnen het vervolg van de dialoog plaatsvindt. Immers, *door* de pressie is het bod profijtelijk geworden, althans zo hoopt de spreker. Via profijtelijkheidsargumentatie toetsen de deelnemers of er een situatie is gecreëerd waarin de premissen van de profijtelijkheidsargumentatie aanvaardbaar zijn voor de ontvanger van het bod.

Pressie komt in gradaties. We onderscheiden twee methoden om de druk te vergroten: door de boodschap dat de rek eruit is en door de inzet te verhogen.

Bij de eerste methode wordt de druk vergroot door de indruk te wekken dat na afwijzing mogelijkwerwijs geen nieuw bod volgt, of zelfs door duidelijk te maken dat dit bod echt het allerlaatste bod is, zodat bij afwijzing de partijen terug bij af zijn.

De tweede methode houdt in dat de beloning voor aanvaarding van het bod wordt verhoogd of de straf voor afwijzing wordt verzwaaard. Het is vaak aannemelijk dat er bij aanvang van de onderhandelingsdialoog overeenstemming bestaat over wat wel en niet behoort tot de inventaris van zaken die tot inzet van de onderhandelingen gemaakt kunnen worden. Bij de aanschaf van een piano is het ruilmiddel de hoeveelheid geld die je bereid bent te betalen en de hoeveelheid geld die de verkoper wenst te ontvangen, en wellicht ook de wijze van betaling. Je kunt de inzet verhogen en de druk opvoeren door meer geld te bieden dan je eerder deed, of door aan te bieden nu contant af te rekenen, in plaats van achteraf giraal.

Je kunt de inzet ook verhogen door de verkoper een bod te doen op basis van een uitbreiding van de inventaris, bijvoorbeeld door aan te bieden om in ruil voor een beperking van de prijs van de piano tot 5000 Euro nu en dan in de winkel piano te komen spelen. Na een dergelijk bod moet de tegenpartij de mogelijkheid hebben om het voorstel tot uitbreiding van de inventaris te aanvaarden ('Okay, ik wil best praten over uw aanbod') dan wel af te wijzen ('Nee, wij nemen alleen geld aan'). Bij aanvaarding heeft deze partij nog steeds de keuze om het specifieke bod te aanvaarden dan wel af te wijzen om er meer uit te slepen.

Je zou de druk ook kunnen opvoeren door de beloning te verhogen op basis van een uitbreiding van de inventaris waarvan niet verwacht kan worden dat die in alle openheid door de andere partij aanvaard kan worden. Dan spreken we van 'omkoperij.' Stel, je biedt de eigenares van de zaak niet alleen 5000 in plaats van 6000 Euro maar belooft ook om als gemeenteraadslid in te stemmen

met een voorstel voor extra parkeerplaatsen direct naast haar muziekwinkel om zo haar verzet tegen de lagere prijs te breken. Ook omkoperij verhoogt de druk.

Ten slotte kan de druk worden opgevoerd door, in plaats van de beloning te verhogen, de sanctie te verzwaren, en te dreigen op een manier die verder gaat dan de ontvanger hoefde te verwachten bij aanvang van de onderhandelingsdialoog, bijvoorbeeld door te dreigen met het zwart maken van de zaak op sociale media mocht je de piano niet meekrijgen voor de door jou geboden prijs. In een dergelijk geval spreken we van ‘chanteren.’ Chanteren verhoogt, net als omkoperij, de druk.

In Voorbeeld 3 is de rek eruit.

Er werd geroepen: Wij zijn niet te stoppen, klimaatverandering wel! En zo is het. Dit is nog maar het begin. De klimaatparade in Amsterdam vandaag (...) was een van de acties in de snel groeiende beweging tegen klimaatverandering (...) Het gaat niet om wat mooie bomen of een leuk bos om in het weekend te gaan wandelen. Het gaat om overleven. In 30 jaar met CO₂ uitstoot naar NUL om enige kans te maken op een leefbare planeet. Wij accepteren geen concessies, geen geleuter. (...) Het is nu op, ons geduld. Het gaat niet om auto's, vliegtuigen, televisies en smartphones. Het gaat om zeespiegelstijging, voedselproductie en overleven. (Wij Stoppen Kool, 2015)

Voorbeeld 3 Wij zijn niet te stoppen

Dit fragment duidt op een bod: ‘Wij zijn bereid onze acties te beëindigen maar alleen als jullie de klimaatverandering een halt toeroepen.’ De druk wordt opgevoerd door aan te geven er verder geen concessies geaccepteerd kunnen worden, zodat dit bod kennelijk ook het laatste bod is.

In Voorbeeld 4 is wellicht sprake van omkoperij.

Staatssecretaris Henk Bleker ontkent dat zijn ambtenaren de bekende natuurbeschermers Jaap Dirkmaat onder druk hebben gezet om de juridische strijd tegen zijn beleid te staken. (...) Maar Dirkmaat beschuldigt Blekers ambtenaren wel degelijk van zulke methodes. Hij zegt na werktijd telefoontjes van ambtenaren te hebben gehad. Deze beloofden hem bijstand in Brussel als hij het gevecht op zou geven. Dirkmaat: ‘Zij vertelden dat Bleker mij steun in Europa zou verlenen voor mijn Vereniging Nederlands Cultuurlandschap. Maar dan moest ik wel afzien van alle juridische stappen tegen hem én de kritische brief intrekken die ik hem medio maart over zijn natuurbeleid stuurde.’ Dirkmaat diende in Brussel een formele klacht in tegen de Nederlandse staat omdat hij vindt dat Blekers bezuinigingen indruisen tegen internationale verdragen. (Nieuwsdienst, 2011)

Voorbeeld 4 Telefoontjes van ambtenaren

Dit kan gezien worden als een voorbeeld van een soort omkoperij: Blekers ambtenaren bieden een beloning aan om Dirkmaat te verlokken een tegenprestatie te leveren die buiten de inventaris van uitruilbare kwesties valt en waarvan niet verwacht kan worden dat Dirkmaat haar alsnog in alle openbaarheid zou kunnen opnemen in deze inventaris.

Volgens onze analyse gaat van een bod altijd enige pressie uit, aangezien een bod verbonden is met een dreigement en met een poging tot verlokken, maar kan de mate van pressie variëren. Hoe dan pressie te beoordelen?

De beoordeling van pressiemiddelen

Vanuit het gezichtspunt van overtuigingsdialoog is het uitoefenen van pressie irrelevant, omdat de pressie niets bijdraagt aan het redelijk oplossen van een verschil van mening. Voor onderhandelingsdialogen vormen dreigen met sancties en verlokken met beloningen echter een *conditio sine qua non*. Zulke tactieken moeten dus beschouwd worden als *prima facie* legitiem in een dergelijke context.

Ook het vergroten van pressie door de indruk te wekken dat je je laatste bod nadert, of zelfs al hebt bereikt – “de rek is eruit” – behoort tot de toelaatbare strategieën van het onderhandelen. Immers, onderhandelen vindt vaak plaats onder tijdsdruk zodat er op een goed moment een einde moet komen aan de dialectiek van bod en tegenbod. En zelfs zonder tijdsdruk moet ieder van de partijen de vrijheid hebben om een scenario zonder akkoord te verkiezen boven een slecht akkoord, en die boodschap aan de andere partij over te dragen.

Het uitbreiden van de inventaris van ruilmiddelen moet eveneens beschouwd worden als een legitiem soort strategie, en wordt in de onderhandelingsliteratuur doorgaans aanbevolen omdat het ‘integratief onderhandelen’ faciliteert. Er is sprake van ‘distributief onderhandelen’ als de partijen gericht zijn op de voor hen gunstigste verdeling van een gegeven goed (Raiffa et al., 2002, p. 97). Bij integratief onderhandelen daarentegen proberen de deelnemers een oplossing te bereiken die beiden tot maximaal voordeel strekt door eerst na te gaan of een vergroting van de inventaris van ruilmiddelen een meer optimale uitruil mogelijk maakt (Raiffa et al., 2002, p. 191). Streven naar vergroting van de inventaris is echter een enigszins risikante strategie, aangezien het initiatief om andere zaken te betrekken bij de onderhandeling ook gezien kan worden als een verwerpelijke vorm van koehandel, of zelfs als chantage of omkoperij.

Toch zijn niet alle gradaties waarmee pressie uitgeoefend kan worden in overeenstemming met het doel van een onderhandelingsdialoog. Vanuit een normatief gezichtspunt worden de partijen in een dialoog van dit type verondersteld redelijke en legitieme middelen te gebruiken om daarmee een compromis te bereiken dat ze vrijwillig onderschrijven. Daarom vallen chantage en omkoperij buiten de orde. Het zijn instrumenten om de tegenpartij te overdonderen, zodat bij gebruik van deze middelen de redelijkheid en de vrijwilligheid van een eventueel akkoord in gevaar zijn. Wat echter in een bepaalde situatie als chantage of omkoperij geldt, is voor verschillende soorten onderhandeling verschillend. Wellicht probeerde men Dirkmaat in Voorbeeld 4 (*Telefoontjes van ambtenaren*) om te kopen, maar is een soortgelijke uitruil van diensten heel gepast in een bepaald soort zakelijke onderhandeling. Zo is het dreigement om tegoeden te bevriezen ongepast in een bankcliëntrelatie, maar past het wel in het bestek van een diplomatieke onderhandeling.

Omkoperij en chantage vormen daarmee overtredingen van normen die gelden in onderhandelingsdialogen. We karakteriseren beide echter niet als “drogreden” aangezien we die term willen reserveren voor overtredingen van regels voor overtuigingsdialogen. Drogredenen zijn daarmee obstakels voor het inhoudelijk oplossen van meningsverschillen, terwijl omkoperij en chantage, zoals wij die termen hier gebruiken, voorbeelden zijn van obstakels voor het bereiken van een onderhandelingsuitkomst dat als een vrijwillig aangegaan compromis kan gelden.

Het is niet uitgesloten dat twee partijen met elkaar zowel in een overtuigingsdialoog als in een onderhandelingsdialoog verweekeld zijn zonder dat ze het verloop van de dialoog van het ene type van invloed laten zijn op het verloop van die van het andere type. Dat is misschien zelfs mogelijk als beide dialogen met elkaar verweven zijn zoals uiteengezet in de vorige paragraaf. Maar dan valt toch eerder te verwachten dat het verloop de ene dialoog het verloop van de andere beïnvloedt. Zo merkten we in de vorige paragraaf al op dat het verloop van een overtuigingsdialoog de keuzeruimte in een onderhandelingsdialoog kan beïnvloeden en dat pressie in een onderhandelingsdialoog invloed kan hebben op het verloop van een overtuigingsdialoog.

Het effect van pressie in een onderhandelingsdialoog die verweven is met een overtuigingsdialoog kan zijn dat de onder druk staande partij afziet van haar streven om haar positie in de overtuigingsdialoog verder uit te werken en ertoe overgaat standpunten, kritische overwegingen of argumenten in te slikken. Bij de boodschap dat een andere partij op een bepaald punt in een overtuigingsdialoog beter zijn mond kan houden, om

op die wijze de sanctie te ontlopen of juist de beloning in de onderhandelingsdialoog te kunnen innen, wordt de normale gang van zaken in de overtuigingsdialoog gehinderd. Hoe meer de druk wordt opgevoerd, des te meer de pressie de soort van samenwerking verstoort die is vereist voor een overtuigingsdialoog; uiteindelijk zou dit ertoe kunnen leiden dat een onder druk staande partij als gevolg van wantrouwen, angst of ergernis niet langer in staat of bereid is deel te nemen aan de overtuigingsdialoog.

Onderhandelingsdialogen laten een zekere mate van pressie toe (maar de soort en mate van toelaatbare pressie is wel afhankelijk van de soort van onderhandelingsdialoog). Overtuigingsdialogen laten geen pressie toe. Dus zodra een partij in een maatschappelijke discussie druk uitoefent in een overtuigingsdialoog, maar ook wanneer zij dat doet in een daarmee verweven onderhandelingsdialoog, en daarmee, bedoeld of onbedoeld, het de andere partij moeilijk of zelfs onmogelijk maakt om haar standpunt uit te dragen of zich kritisch te manifesteren in de overtuigingsdialoog, dan spreken we van een drogreden met betrekking tot de overtuigingsdialoog: een drogreden *ad baculum* als de pressie meer het karakter heeft van een dreigement; een drogreden *ad carotam* als de beloning meer op de voorgrond staat.⁵

In het geval de druk wordt uitgeoefend in een onderhandelingsdialoog maar invloed heeft op een overtuigingsdialoog spreken we van een geïnduceerde drogreden. Dat zulke effecten zich voordoen valt niet altijd te vermijden, maar ook in dat geval is niet uitgesloten dat de maatschappelijke discussie als geheel tot een redelijke uitkomst leidt. Bovendien kan een partij goede redenen hebben om druk uit te oefenen, bijvoorbeeld om tegenspel te bieden aan druk die de andere partij uitoefent, en daarmee deze partij ertoe te brengen zich wat redelijker op te stellen (Jacobs, 2009; Van Laar & Krabbe, 2016a, 2016b).

Conclusie

Het gebruik van pressie is soms noodzakelijk. Wanneer er onderhandeld wordt is een zekere mate van pressie onvermijdelijk, maar daarmee zijn, zelfs in dat geval, nog niet alle soorten van pressie legitiem. Het betekent ook niet dat er geen drogredenen *ad baculum* en *ad carotam* bestaan: deze kunnen zich wel degelijk voordoen in de overtuigingsdialogen die ook deel uitmaken van een maatschappelijke discussie. Door de verwevenheid van onderhandelings- en overtuigingsdialogen binnen een maatschappelijke discussie kan de legitieme pressie die in de eerste soort dialogen wordt

uitgeoefend deze drogredenen induceren in de dialogen van de tweede soort. Vanuit het normatieve gezichtspunt van onderhandelingsdialogen mag er dan niet veel aan de hand zijn, vanuit dat van overtuigingsdialogen is er dan toch iets mis wat de argumentatie betreft.

Het is een taak voor de argumentatietheorie om ook bij dit soort situaties, waarin discussiebijdragen vanuit één normatief gezichtspunt normaal zijn en vanuit een ander abnormaal de discussianten te voorzien van de nodige middelen van verweer.

We concluderen dat pressiemiddelen, of het nu om dreigementen of verlokkingen gaat, een redelijke uitkomst van de maatschappelijke discussie kunnen bevorderen zolang het om soorten van pressie gaat die eigen zijn aan de soort onderhandelingsdialogen waarin de pressie wordt uitgeoefend, maar anderzijds dat deze middelen een redelijke uitkomst in de weg kunnen staan, namelijk wanneer pressie – al dan niet via onderhandelingsdialogen – een effect heeft op het verloop van een overtuigingsdialoog die ook deel uitmaakt van de maatschappelijke discussie.

Noten

1. Andere eigenschappen van een maatschappelijke discussie zijn: dat deze doorgaans diverse, verwante kwesties betreft, dat er doorgaans meer dan twee partijen zijn, dat ze een zekere duur heeft, dat ze zich voor een belangrijk deel afspeelt in de massa media, en dat er doorgaans voor ten minste enkele betrokken veel op het spel staat, zoals hun inkomen, hun manier van leven of hun identiteit (Van Laar & Krabbe, 2016a, 2016b).
2. De keuze voor deze maatschappelijke discussie is mede bepaald door ons streven een bijdrage te leveren aan het COST Action project "European network for argumentation and public policy analysis" (CA17132), dat zich richt op de studie van de klimaat- en energiediscussie in de EU.
3. De argumentaties die tot uiting komen in een uitwisseling van bod en tegenbod kunnen gezien worden als een dialogische variant van een argumentative pattern dat een bepaalde institutioneel ingebedde communicatieve activiteit kan kenmerken. Volgens van Eemeren wordt een argumentative pattern 'characterized by a constellation of argumentative moves in which, in order to deal with a particular kind of difference of opinion, in defence of a particular type of standpoint a particular argument scheme or combination of argument schemes is used in a particular kind of argumentation structure' (Van Eemeren, 2016, p. 14).
4. Onze analyse van een bod is in overeenstemming met Ihnen Jory's analyse van de taalhandeling van het doen van een voorwaardelijk bod (Ihnen Jory, 2016).
5. Theoretisch zijn beide drogredenen tot elkaar te herleiden maar doorgaans is het de ontvanger van een bod wel duidelijk of er een wortel te vergeven valt, of dat er een stok achter de deur zit.

Bibliografie

- Amgoud, L., & Prade, H. (2006). Formal handling of threats and rewards in a negotiation dialogue. In S. Parsons, N. Maudet, P. Moraitis, & I. Rahwan (Eds.), *Argumentation in Multi-agent Systems: Second International Workshop, ArgMAS 2005, Utrecht, The Netherlands, July 26, 2005: Revised selected and invited papers* (pp. 88-103). Berlin: Springer.
- Du Pré, R. (2016, 26 oktober). 'Nederland moet voortouw nemen bij klimaatdoelen.' *Volkskrant* (rubriek Ten Eerste), p. 4.
- D66 (2016). D66 Verkiezingsprogramma 2017-2021. Geraadpleegd op 2 januari 2018, <https://d66.nl/content/uploads/sites/2/2016/12/Definitieve-VKP.pdf>
- Eemeren, F.H. van, (2016). Identifying argumentative patterns: A vital step in the development of pragma-dialectics. *Argumentation* 30, 1-23.
- Eemeren, F.H. van, & Grootendorst, R. (2000). *Kritische discussie*. Meppel: Boom.
- Eemeren, F.H. van, & Grootendorst, R. (2004). *A Systematic Theory of Argumentation*. Cambridge: Cambridge University Press.
- Ihnen Jory, C. (2016). Negotiation and deliberation: Grasping the difference. *Argumentation* 30, 145-165.
- Jacobs, S. (2009). Nonfallacious rhetorical design in argumentation. In Frans H. van Eemeren and Bart Garssen (Eds.), *Pondering on Problems of Argumentation: Twenty Essays on Theoretical Issues* (pp. 55-78). Houten: Springer.
- Laar, J.A van., & Krabbe, E.C.W. (2016a). Eerlijke en oneerlijke strategieën in maatschappelijke discussies. In D. Van De Mierop, L. Buysse, R. Coesemans, & P. Gillaerts (Red.), *De macht van de taal: Taalbeheersingsonderzoek in Nederland en Vlaanderen* (pp. 131-143), Leuven: Acco.
- Laar, J.A. van, & Krabbe, E.C.W. (2016b). Fair and unfair strategies in public controversies. *Journal of Argumentation in Context* 5(3), 316-348.
- Laar, J.A. van, & Krabbe, E.C.W. (2016c). Splitting a difference of opinion. In P. Bondy, & L. Benacquista (Eds.), *Argumentation, Objectivity, and Bias: Proceedings of the nth International Conference of the Ontario Society for the Study of Argumentation (OSSA), May 18-21, 2016*. Windsor, ON: OSSA.
- Laar, J.A. van, & Krabbe, E.C.W. (2018). The role of argument in negotiation. *Argumentation*, 32(4), 549-567.
- Nieuwsdienst (2011, 13 mei). Staatssecretaris Bleker ontkent druk op natuurbeschermers. *Algemeen Dagblad*, p. 4
- Raiffa, H., Richardson, J., & Metcalfe, D. (2002). *Negotiation Analysis: The Science and Art of Collaborative Decision Making*. Cambridge, MA: The Belknap Press of Harvard University Press.
- Walton, D.N. (1998). *The New Dialectic: Conversational Contexts of Argument*. Toronto: University of Toronto Press.
- Walton, D.N. (2000). *Scare Tactics: Arguments that Appeal to Fear and Threats*. Dordrecht: Kluwer.
- Walton, D.N. (2014). A dialectical analysis of the *Ad Baculum Fallacy*. *Informal Logic*, 34, 276-310.
- Walton, D.N. & Krabbe, E.C.W. (1995). *Commitment in Dialogue: Basic Concepts of Interpersonal Reasoning*. Albany, NY: State University of New York Press.
- Wij Stoppen Kool (2015). Wij zijn niet te stoppen, klimaatverandering wel! Geraadpleegd op 3 januari 2018, <http://www.wijstoppensteenkool.nl/?p=7628>
- Woods, J. (2004). *The Death of Argument: Fallacies in Agent Based Reasoning*. Dordrecht: Kluwer.

De 'stem van het volk'

Populisme en perspectief

Maarten van Leeuwen en Freek van Vliet
Universiteit Leiden

TVT 41 (1): 181–195

DOI: 10.5117/TVT2019.1.012.LEEU

Abstract

The 'voice of the people': populism and perspective

Recently, a growing body of (political-scientific) literature has focused on the empirical measurement of populism. In such studies, "people-centrism" is one of the most frequently analysed discourse characteristics, i.e. to what extent "the people" are put in the focus of attention in a politician's discourse. In order to measure people-centrism empirically, it is common practice to use the number of references to the electorate as the only indicator. In this contribution, we substantiate that a linguistic-stylistic analysis of "perspective" or "attributed viewpoint" offers additional tools for analysing people-centrality in (populist) political discourse empirically. As a case study, we report on an analysis of series of parliamentary speeches delivered by the Dutch populist politicians Geert Wilders and Rita Verdonk.

Keywords: Stylistic analysis, speech & thought representation, measuring populism, people-centrism, Geert Wilders, Rita Verdonk

Inleiding¹

In (met name politicologisch) onderzoek naar populisme is het afgelopen decennium een groeiende belangstelling voor de vraag hoe populisme empirisch kan worden gemeten (Akkerman, Mudde, & Zaslove, 2014, p. 5; Rooduijn & Pauwels, 2011, p. 1272). Daartoe wordt in toenemende mate geprobeerd op basis van systematische tekstanalyse vast te stellen wat

het populistisch gehalte is van een politicus of politieke partij (zie bijv. Rooduijn, De Lange, & Van der Burg, 2014; Hawkins, 2009). Een van de meest frequent gebruikte maatstaven voor uitspraken over het populistisch gehalte van politici is daarbij wat in de literatuur *people centrality* wordt genoemd (Rooduijn & Akkerman, 2017, p. 194). Deze notie verwijst naar de mate waarin politici 'het volk' ofwel 'de mensen in het land' centraal stellen in hun discours. De gerichtheid op *people centrality* valt goed te verklaren: hoewel er onder politicologen nog geen consensus is over de vraag wat populisme precies is (Moffitt, 2016), wordt algemeen erkend dat het centraal stellen van kiezers een belangrijk onderdeel vormt van een populistisch discours (Rooduijn & Akkerman, 2017, p. 194). Door 'het volk' zeer centraal te stellen, zo is de gedachte, suggereren populistten (meer dan andere politici) dat zij de wil van de gewone man vertolken.²

Om vast te stellen in welke mate politici het volk centraal stellen in hun discours, is het gebruikelijk om het aantal verwijzingen naar kiezers als enige maatstaf te hanteren (zie bijv. Rooduijn & Akkerman, 2017; Oliver & Rahn, 2016; Pauwels, 2014; Jagers & Walgrave, 2007). In dit artikel willen wij laten zien hoe een taalkundig-stilistische analyse van perspectief (in termen van *speech and thought representation*) de analist extra handvatten biedt om de mate van *people centrality* te meten. Daartoe richten we op ons op een casus: de mate waarin en manier waarop Geert Wilders (PVV) en Rita Verdonk (Trots op Nederland) kiezers een eigen perspectief toeschrijven in een serie parlementaire toespraken in de periode 2008-2010.

Dit artikel is als volgt opgebouwd. In de volgende paragraaf gaan we eerst kort in op het idee dat niet alleen het aantal verwijzingen naar kiezers als maatstaf kan worden gehanteerd voor de mate waarin kiezers centraal worden gesteld in het discours van een politicus, maar dat dat evenzeer geldt voor de mate waarin (en manier waarop) kiezers een eigen perspectief krijgen toegeschreven. Daarna behandelen we onze casus: we schetsen de aanleiding, het geanalyseerde corpus, onze methode en de resultaten. In de conclusie vatten we de belangrijkste uitkomsten van onze gevalsstudie samen.

Kiezers centraal stellen en perspectief

Het idee dat de mate van *people centrality* samenhangt met de frequentie waarin politici naar kiezers verwijzen, wordt in de invloedrijke studie van Jagers & Walgrave (2007) als volgt gemotiveerd:

Political actors (...) frequently use words such as '(the) people', '(the) public', '(the) citizen(s)', '(the) voter(s)', '(the) taxpayer(s)', '(the) resident(s)', '(the) consumer(s)' and '(the) population'. By referring to the people, a political actor claims that he or she cares about the people's concerns, that he or she primarily wants to defend the interests of the people, that he or she is not alienated from the public but knows what the people really want. The implicit (...) motto is: 'I listen to you because I talk about you.' (Jagers & Walgrave, 2007, p. 323)

Kiezers centraal stellen is echter niet alleen een kwestie van frequent naar kiezers verwijzen. Een andere factor is de mate waarin politici kiezers een eigen 'gezichtspunt' of perspectief toeschrijven in hun discours. Door niet alleen *over* kiezers te spreken, maar ook weer te geven wat kiezers zeggen en denken, wekken politici de indruk dat kiezers in de discussie worden betrokken: kiezers krijgen als het ware een eigen 'stem' in het debat (vgl. Van Leeuwen, 2015, p. 125). In termen van Clarks (1996) participantenrollen worden zij niet geconstrueerd als 'overhearers' of 'bystanders', maar als 'side participants' die daadwerkelijk in de discussie participeren.

Onderzoek naar *speech and thought representation* heeft laten zien dat schrijvers en sprekers op verschillende manieren kunnen weergeven wat andere mensen zeggen en denken. Vier frequent terugkerende categorieën in de literatuur zijn het gebruik van de directe rede, de vrije indirecte rede, de indirecte rede en 'impliciet perspectief' (zie bijv. Sanders, 2009; 1994), waarvan fragmenten (1)-(4) achtereenvolgens (verzonnen) voorbeelden zijn.

- (1) Henk en Ingrid zeiden tegen mij: *'We zijn grote tegenstanders van belastingverhoging.'*
- (2) Henk en Ingrid zeiden het zelf. *Ja, ze waren grote tegenstanders van belastingverhoging.*
- (3) Henk en Ingrid zeiden *dat ze grote tegenstanders van belastingverhoging waren.*
- (4) Henk en Ingrid *willen absoluut geen belastingverhoging.*

Sanders (2009; 1994) betoogt dat deze vormen van perspectivering onderling verschillen in de mate waarin het perspectief van het opgevoerde 'personage' (in (1)-(4): 'Henk en Ingrid') op de voorgrond wordt gebracht. Meer specifiek betoogt ze op basis van cognitief-linguïstische inzichten dat de perspectiefcategorieën waar (1)-(4) voorbeelden van zijn, op een schaal kunnen worden geplaatst van relatief sterk geperspectiverd ((1)) naar relatief zwak geperspectiverd ((4)). Bij de directe rede ((1)) wordt het opgevoerde personage maximaal gepresenteerd als verantwoordelijk voor de

inhoud en vorm van de uiting, terwijl bij een impliciet perspectief ((4)) die verantwoordelijkheid voor vorm en inhoud grotendeels bij de verteller ligt. Impliciet perspectief komt tot uitdrukking in werkwoorden en werkwoordelijke constructies van perceptie (*zien, ontdekken*), cognitie (*weten, willen, dromen, aarzelen*), emotie (*vrezen, verheugen*) en modaliteit (*moeten, kunnen*) (Sanders, 2009, p. 3; Sanders, 1994, pp. 55-63). De vrije indirecte rede ((2)) en de directe rede ((3)) liggen tussen deze uitersten in, waarbij de vrije indirecte rede als sterker geperspectieerd wordt beschouwd dan de indirecte rede, omdat bij de vrije indirecte rede de verantwoordelijkheid voor inhoud en vorm van de uiting tussen verteller en opgevoerd personage gedeeld wordt (de uiting is ambigu), terwijl bij de indirecte rede een groot deel van de vorm van de uiting bij de verteller ligt (Sanders, 2009, p. 2; Sanders, 1994, pp. 30-81). Wat de verschillende perspectiefcategorieën echter met elkaar gemeen hebben, is dat het opgevoerde personage (in meer of mindere mate) een eigen 'stem' krijgt in een tekst. Daarvan kunnen politici strategisch gebruikmaken: door kiezers in meer of mindere mate een eigen 'perspectief' toe te schrijven, wekken ze in meer of mindere mate de suggestie dat kiezers 'meedoen' in de discussie.³

Casus: Verdonk versus Wilders

Aanleiding

Om te illustreren hoe een taalkundig-stilistische analyse van perspectief kan bijdragen aan het meten van de mate waarin politici kiezers centraal stellen in hun discours, analyseren we in het vervolg van dit artikel een serie parlementaire toespraken uit de periode 2008-2010 van twee Nederlandse politici die door politicologen als 'populistisch' zijn bestempeld: Rita Verdonk (Trots op Nederland, ToN) en Geert Wilders (Partij Voor de Vrijheid, PVV). Aanleiding vormt de studie van Vossen (2009), die beide politici langs een populistische meetlat heeft gelegd. Aan de hand van zeven kenmerken van populisme (twee 'basingrediënten' en vijf 'smaakversterkers') concludeert hij, op basis van een analyse van allerlei bronnenmateriaal uit de periode 2003-2009, dat de populismegraad bij Rita Verdonk hoger ligt dan bij Geert Wilders. Vossen (2009) betoogt onder meer dat Verdonk in de onderzochte periode meer dan Wilders aan 'volksverheerlijking' doet.

Vossens kwalitatieve en 'wellicht soms wat impressionistische analyse' (Vossen, 2009, p. 456) is grotendeels gebaseerd op bronnenmateriaal van beide politici van *buiten* de Tweede Kamer (interviews, portretten,

reportages, partijdocumenten en een aantal toespraken).⁴ Omdat de Tweede Kamer voor politici een belangrijk platform is om 'de stem van het volk' te representeren, is Vossens studie voor ons aanleiding systematisch te onderzoeken in hoeverre beide politici *in de Tweede Kamer* kiezers centraal stellen in hun discours. Anders dan gebruikelijk (zie bijv. Rooduijn & Akkerman, 2017; Oliver & Rahn, 2016; Pauwels, 2014; Jagers & Walgrave, 2007) onderzoeken we daartoe niet alleen systematisch de mate waarin beide politici in hun parlementaire speeches naar kiezers verwijzen, maar ook in hoeverre en op wat voor manier beide politici kiezers een eigen perspectief toeschrijven. Geldt ook voor Verdonks parlementaire bijdragen dat kiezers een centralere rol krijgen toebedeeld dan bij Wilders – hetgeen op basis van Vossens analyse verwacht zou kunnen worden? Verwijst Verdonk in haar parlementaire speeches vaker naar kiezers, en kent zij kiezers in sterkere mate dan Wilders een eigen perspectief toe?

Corpusamenstelling

Voor de analyse hebben we ons gericht op speeches van beide politici in de Tweede Kamer zoals die in de Handelingen van de Tweede Kamer zijn terug te vinden.⁵ Meer specifiek hebben we in de periode 15 september 2007 (de dag waarop Rita Verdonk uit de VVD werd gezet en zelfstandig Kamerlid werd) – 16 juni 2010 (de dag waarop Verdonk de politiek verliet)

Tabel 1 Overzicht geanalyseerde speeches en het aantal woorden per toespraak

	Datum speech	Tijdens welk debat?	Aantal woorden Wilders	Aantal woorden Verdonk ⁷
1	22-05-2008	Verantwoordingsdebat 2008	2.080	1.145
2	05-06-2008	Spoeddebat over de beveiliging van mevrouw Verdonk	533	823
3	17-09-2008	Algemene Politieke Beschouwingen 2008	2.895	830
4	30-09-2008	Debat over de situatie op de internationale financiële markten	769	436
5	18-11-2008	Spoeddebat over de kredietcrisis	726	690
6	18-11-2008	Spoeddebat over Ella Vogelaar	479	395
7	23-04-2009	Debat over de JSF	1.002	363
8	28-05-2009	Verantwoordingsdebat 2009	2.149	870
9	16-09-2009	Algemene Politieke Beschouwingen 2009	3.150	955
10	18-02-2010	Debat over Uruzgan	551	359
11	20-05-2010	Verantwoordingsdebat 2010	1.354	866
		Totaal	15.688	7.732

alle debatten geïdentificeerd waarin zowel Verdonk als Wilders een toespraak hield.⁶ In deze periode voerden beide politici als fractievoorzitter oppositie tegen het kabinet-Balkenende IV. Een overzicht van de geanalyseerde toespraken is te vinden in Tabel 1.

Analysemethode

In de corpusanalyse zijn eerst alle verwijzingen naar kiezers geteld, waarna is vastgesteld in welke mate kiezers een eigen perspectief krijgen toegeschreven. Op beide analyseonderdelen gaan we hieronder nader in. Beide deelanalyses zijn verricht door de tweede auteur, op basis van de hieronder beschreven criteria. De afwezigheid van een interbeoordelaarsbetrouwbaarheids-berekening kan daarbij vragen oproepen over de betrouwbaarheid van het onderzoek: werken met één codeur heeft immers als nadeel dat dit kan leiden tot idiosyncrasieën in de analyse. De vraag is echter wel wat de invloed van dergelijke idiosyncrasieën is op de uitkomst van het onderzoek: in de gerapporteerde stilistische analyses staat de *vergelijking* tussen Wilders en Verdonk centraal. Aangenomen mag worden dat eventuele idiosyncrasieën van de codeur een systematisch karakter zullen hebben en zich dus bij beide deelcorpora zullen voordoen; dit betekent dat ze de *vergelijking* tussen Wilders en Verdonk, waar het in de analyse om gaat, in principe niet vertroebelen (zie ook Van Leeuwen, 2015, pp. 35-36 en de referenties aldaar).

Verwijzingen naar de kiezer

In navolging van Van Leeuwen (2015, p. 43) zijn verschillende categorieën als een verwijzing naar de kiezer opgevat. Meegenomen zijn verwijzingen naar individuen (zie (5)-(6)) of groepen mensen in de samenleving (zie (7)-(8)),⁸ verwijzingen naar 'Nederland', 'de samenleving' en 'de maatschappij' wanneer deze metonymisch voor de inwoners van Nederland worden gebruikt (zoals in (9)), onbepaalde voornaamwoorden als 'iedereen' en 'niemand' (zie (10)) en persoonlijke voornaamwoorden als 'hen' en 'zij' wanneer deze expliciet te koppelen zijn aan een 'kiezer' eerder in de tekst (zie (11)).

- (5) *Henk en Ingrid* willen dolgraag dat de studiefinanciering voor hun kinderen wordt gehandhaafd. [W₁]⁹
- (6) Als er geld te kort is, is er altijd nog *de burger* om aan te plukken. [V₁]
- (7) Dat betekent dat *starters op de woningmarkt* nog veel langer zullen moeten wachten (...)? [V₃]
- (8) Vertelt u dat eens aan *die agenten in Gouda* (...) [W₃]

- (9) *Nederland* heeft zelden zo'n zootje ongeregeld bij elkaar gezien. [W10]
- (10) Er gaat 4,7 miljard Nederlands belastinggeld naar Griekenland en *ieder-een* weet dat dit geld nooit meer terugkomt. [V11]
- (11) *Hoogbejaarden* krijgen doorzitwonden omdat *zij* dagen moeten doorbrengen in aftandse rolstoelen. Thuiszorghulpen worden *hen* afgepakt. [W1]

Om het aantal verwijzingen naar kiezers te tellen, zijn de speeches van beide politici eerst met de tekstsegmentatieprocedure van Pander Maat (2002, pp. 37-43) opgedeeld in *taaluitingen*.¹⁰ Vervolgens is per taaluiting vastgesteld of hierin wel of niet naar kiezers wordt verwezen. Fragment (12) bevat bijvoorbeeld twee taaluitingen, waarbij elk van deze uitingen één verwijzing naar de kiezer bevat. In fragment (13) is sprake van één taaluiting; de drie verwijzingen naar kiezers daarin zijn eveneens als één meegerekend.¹¹

- (12) [De kredietverlening aan *ondernemers* is zeer moeizaam geworden] [en *bankmedewerkers* worden met duizenden tegelijk naar huis gestuurd.] [V7]
- (13) Dit kabinet mag onze *ouderen*, *gehandicapten* en *ernstig zieken* niet in de kou laten staan. [W3]

Perspectief

Om te meten in hoeverre Wilders en Verdonk kiezers een eigen perspectief toeschrijven, is voor alle taaluitingen waarin Wilders en Verdonk naar kiezers verwijzen geanalyseerd of in deze taaluitingen kiezers een eigen perspectief krijgen toegeschreven, en zo ja, op wat voor manier (is er sprake van directe rede, vrije indirecte rede, indirecte rede of impliciet perspectief?).¹² Hiertoe is gebruikgemaakt van een licht aangepaste versie van het analysemodel van Sanders (1994, p. 67), dat hieronder is weergegeven.¹³ Sanders' analysemodel stelt de analist in staat om aan de hand van drie noties ('propositionele inhoud', 'referentieel centrum' en 'deiktisch centrum') stapsgewijs voor iedere taaluiting vast te stellen of een opgevoerd personage (hier: 'de kiezer') een eigen perspectief krijgt toegeschreven, en zo ja, met welk van de vier genoemde perspectiefformen (zie Sanders, 1994, pp. 30-81 voor details). Bij de analyse van impliciet perspectief zijn daarbij alle in de tweede paragraaf genoemde vormen van impliciet perspectief meegenomen.

Model voor de analyse van perspectief (gebaseerd op Sanders, 1994, p. 67)

Stap 1: Bepaal wie verantwoordelijk is voor de propositionele inhoud (P) van de taaluiting.

- De politicus zelf? → ga naar stap 4
- De kiezer? → ingebed perspectief → ga naar stap 2

Stap 2: Bepaal hoe de ‘stem’ van de kiezer gerepresenteerd wordt door de politicus.

- Zijn de propositionele inhoud (P), het *referentiële centrum* (R) én het *deiktisch centrum* (D) allemaal afkomstig van de kiezer? → directe rede
- Is het *referentiële centrum* niet afkomstig van de kiezer? → ga naar stap 3

Stap 3: Bepaal of alleen het *referentieel centrum* (R) bij de politicus ligt, of zowel het *referentiële* als het *deiktische centrum* (D).

- *Referentieel centrum* ligt bij politicus en *deiktisch centrum* bij de kiezer? → vrije indirecte rede
- *Referentieel centrum* én *deiktisch centrum* ligt bij politicus? → indirecte rede

Stap 4: Impliciet perspectief

Bepaal of de politicus wel verwijst naar het bewustzijn van het personage.

- Ja? → impliciet perspectief
- Nee? → geen perspectivering

Resultaten

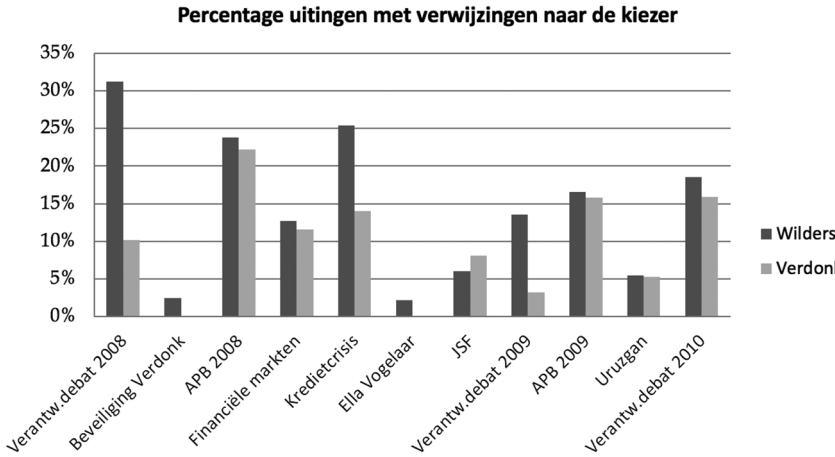
Verwijzingen naar de kiezer

Op basis van Vossens (2009) conclusie dat Rita Verdonk hoger scoort op de populistische meetlat dan Geert Wilders, was onze verwachting dat Verdonk in haar parlementaire speeches vaker dan Wilders naar kiezers verwijst. Deze hypothese blijkt niet te kloppen (zie Tabel 2) – integendeel: Wilders verwijst significant vaker naar de kiezer dan Verdonk ($p < 0,0001$).¹⁴

Omdat het aantal uitingen per speech laag is, is het lastig om uitspraken te doen over de verschillen per afzonderlijk debat. Toch is het opvallend dat Wilders in 10 van de 11 onderzochte debatten relatief gezien meer naar de kiezer verwijst dan Verdonk – er lijkt dus sprake van een consistent patroon. Daarnaast valt op dat de mate waarin beide politici naar kiezers verwijzen per debat nogal verschilt, waarbij Wilders en Verdonk grotendeels overeenkomen wat betreft speeches waarin zij relatief vaak en relatief weinig naar kiezers verwijzen (vgl. Figuur 1). De speeches tijdens de Algemene

Tabel 2 Aantal uitingen waarin Wilders en Verdonk naar de kiezer verwijzen: absolute aantallen en percentages van het totaal

		Rita Verdonk			Geert Wilders		
		Uitingen totaal	Uitingen met kiezer	% van totaal	Uitingen totaal	Uitingen met kiezer	% van totaal
1	Verantwoordingsdebat 2008	98	10	10,2%	189	59	31,2%
2	Beveiliging Verdonk	60	0	0,0%	40	1	2,5%
3	APB 2008	63	14	22,2%	265	63	23,8%
4	Situatie financiële markten	43	5	11,6%	71	9	12,7%
5	Kredietcrisis	43	6	14,0%	67	17	25,4%
6	Ella Vogelaar	34	0	0,0%	46	1	2,2%
7	JSF	37	3	8,1%	100	6	6,0%
8	Verantwoordingsdebat 2009	93	3	3,2%	184	25	13,6%
9	APB 2009 ¹⁵	95	15	15,8%	295	49	16,6%
10	Uruzgan	38	2	5,3%	55	3	5,5%
11	Verantwoordingsdebat 2010	82	13	15,9%	124	23	18,5%
Totaal		686	71	10,3%	1436	256	17,8%



Figuur 1 Percentage uitingen waarin Wilders en Verdonk naar de kiezer verwijzen, uitgesplitst per debat

Beschouwingen en de Verantwoordingsdebatten behoren bij beide politici tot de toespraken waarin het meest naar kiezers wordt verwezen, samen met het debat over de kredietcrisis.¹⁶ In debatten over meer ‘technische’

beleidskwesties, zoals over de beveiliging van Verdonk of het aftreden van Ella Vogelaar gebruiken zowel Wilders als Verdonk het minst verwijzingen naar kiezers (zie Figuur 1). De mate waarin beide politici naar kiezers verwijzen lijkt dus te worden beïnvloed door het debatonderwerp; mogelijk speelt ook de hoeveelheid media-aandacht een rol.¹⁷

Perspectief

Op basis van Vossens (2009) conclusies was onze verwachting dat Verdonk kiezers in sterkere mate dan Wilders een eigen perspectief zou toeschrijven. Ook hier blijkt uit onze resultaten het tegengestelde (vgl. Tabel 3): het percentage uitingen waarin Wilders kiezers perspectief toekent, ligt significant hoger dan bij Verdonk ($p < 0,01$).¹⁸

Tabel 3 Aantal uitingen waarin Wilders en Verdonk de kiezer een perspectief geven: absolute aantallen en percentages van het totaal aantal uitingen

		Rita Verdonk			Geert Wilders		
		Uitingen totaal	Perspec- tief bij kiezer	% van totaal	Uitingen totaal	Perspectief bij kiezer	% van totaal
1	Verantwoor- dingsdebat 2008	98	3	3,1%	189	16	8,5%
2	Beveiliging Verdonk	60	0	0,0%	40	1	2,5%
3	APB 2008	63	4	6,3%	265	24	9,1%
4	Situatie financiële markten	43	3	7,0%	71	4	5,6%
5	Kredietcrisis	43	1	2,3%	67	5	7,5%
6	Ella Vogelaar	34	0	0,0%	46	0	0,0%
7	JSF	37	1	2,7%	100	3	3,0%
8	Verantwoor- dingsdebat 2009	93	0	0,0%	184	11	6,0%
9	APB 2009	95	8	8,4%	295	21	7,1%
10	Uruzgan	38	1	2,6%	55	3	5,5%
11	Verantwoor- dingsdebat 2010	82	3	3,7%	124	10	8,1%
Totaal		686	24	3,5%	1436	98	6,8%

Wat betreft het *type* perspectief dat Wilders en Verdonk gebruiken, is er tussen beide politici minder verschil (zie Tabel 4). Zowel Wilders als Verdonk maken vooral gebruik van impliciet perspectief om kiezers een eigen 'stem' in de discussie te geven, gevolgd door de indirecte rede. Opvallend is echter wel dat Wilders, anders dan Verdonk, ook de directe en vrije indirecte rede

hanteert.¹⁹ Het feit dat Wilders ook relatief sterke perspectiveringstechnieken hanteert, en Verdonk niet, is in lijn met de overige resultaten: Wilders lijkt ook in dit opzicht kiezers een centralere rol toe te kennen in zijn parlementaire speeches dan Verdonk.

Tabel 4 Aantal uitingen waarin Wilders en Verdonk de kiezer een perspectief geven, uitgesplitst naar type perspectief

	Uitingen totaal	Perspectief bij kiezer	Directe rede	Vrije indirecte rede	Indirecte rede	Impliciet perspectief
Verdonk	686	24	0 (0,0%)	0 (0,0%)	4 (20,0%)	20 (83,3%)
Wilders	1436	98	4 (4,3%)	8 (8,2%)	9 (9,2%)	77 (78,6%)

Conclusie

In politicologisch onderzoek naar populisme wordt *people centrality*, d.w.z. de mate waarin kiezers centraal worden gesteld in het discours van een politicus of politieke partij, gezien als een belangrijke indicator voor het populistisch gehalte van die politicus of politieke partij. Om de mate van *people centrality* te meten, wordt de mate waarin politici naar kiezers verwijzen doorgaans als enige maatstaf genomen. In deze bijdrage hebben we willen laten zien hoe een systematische taalkundig-stilistische analyse van perspectief (in termen van *speech and thought representation*) een bijdrage kan leveren aan het nauwkeuriger vaststellen van de mate van *people centrality*. Een analyse van de mate waarin Geert Wilders en Rita Verdonk in hun parlementaire speeches niet alleen naar kiezer verwijzen, maar hun ook een eigen perspectief toeschrijven (en op wat voor manier dan precies), diende hiertoe als casus. Onze gevalsstudie laat zien dat Wilders in zijn parlementaire speeches kiezers niet alleen een centralere rol toekent dan Verdonk door vaker naar kiezers te verwijzen, maar ook door kiezers vaker een eigen perspectief toe te schrijven – inclusief relatief ‘sterke’ vormen van perspectieftoekenning.

De bevindingen van onze systematische taalkundig-stilistische analyse lijken in tegenspraak met de ‘wellicht soms wat impressionistische’ analyse van Vossen (2009, p. 456), die immers concludeerde dat Verdonk hoger scoorde op de populistische meetlat dan Wilders. Een directe vergelijking tussen de hier gerapporteerde uitkomsten en die van Vossen is echter lastig, o.a. vanwege het verschil in onderzocht corpusmateriaal (parlementaire speeches vs. bronnenmateriaal *buiten* de Tweede Kamer) en onderzochte periode (2008-2010 vs. 2003-2009). Het moge duidelijk zijn dat zo’n

directe vergelijking ook niet zozeer het doel was van dit artikel: Vossens studie vormde voor ons vooral een gemotiveerde aanleiding om Wilders' en Verdonks parlementaire speeches als casusmateriaal te selecteren. Aan de hand van dat materiaal hebben wij willen demonstreren hoe een systematische analyse van perspectief een relevante bijdrage kan leveren aan de studie van populisme: door niet alleen te tellen hoe vaak politici naar kiezers verwijzen in hun discours, maar ook de mate waarin en manier waarop kiezers een eigen perspectief krijgen toegeschreven in kaart te brengen, ontstaat een preciezer beeld van de mate waarin politici kiezers centraal stellen in hun discours. Een systematische perspectiefanalyse biedt de analist, met andere woorden, een extra 'tool' om de mate van *people centrality* in politiek discours te meten.

Noten

1. Wij danken twee anonieme reviewers voor hun waardevolle commentaar op een eerdere versie van dit artikel.
2. Binnen de politicologie is er discussie over de vraag of 'populisme' als een absoluut of relatief verschijnsel moet worden gezien (Rooduijn & Akkerman, 2017, pp. 194-195). In de eerste opvatting wordt aangenomen dat politici als 'populist' of 'niet-populist' kunnen worden geclassificeerd. In de tweede opvatting is populisme een gradueel verschijnsel: politici kunnen in 'meerdere' of 'mindere' mate populistische trekken vertonen. De meeste studies die zich richten op het empirisch meten van populisme hanteren deze tweede opvatting; dat geldt ook voor dit artikel.
3. Zoals uit bovenstaande blijkt, wordt de notie 'perspectief' in dit artikel ingevuld in termen van *speech and thought representation*. De notie 'perspectief' wordt daarnaast ook verbonden aan bijvoorbeeld grammaticale positie: in de cognitieve taalkunde wordt ervan uitgegaan dat *het subject* van een zin de 'lens' of 'camerapositie' is van waaruit de informatie die in een zin beschreven staat, primair gepresenteerd wordt (Van Krieken, Sanders, & Hoeken, 2015, p. 222 en de referenties aldaar). Voor het meten van *people centrality* is het zo beschouwd ook interessant om vast te stellen in hoeverre verwijzingen naar kiezers in subjectpositie staan. Zie Van Leeuwen (2015, pp. 112-120) voor een nadere uitwerking en illustratie van dit idee.
4. Vossens (2009) corpus bevat ook een beperkt aantal parlementaire bijdragen, maar deze worden niet systematisch geanalyseerd en in zijn analyse primair aangehaald om de aanwezigheid van *andere* kenmerken van populisme te illustreren.
5. De analyses zijn dus gericht op de *voorbereide bijdragen* aan parlementaire debatten; om de omvang van het corpus binnen de perken te houden zijn interrupties, reacties daarop en ingediende moties buiten beschouwing gelaten.
6. Via de website <https://zoek.officielebekendmakingen.nl/>. Het onderzoek heeft zich dus gericht op de volledige periode waarin Verdonk als zelfstandig Kamerlid optrad (2007-2010). Omdat er in 2007 geen debatten bleken te zijn waarin zowel Verdonk als Wilders spraken, beslaat de feitelijke analyse de periode mei 2008-mei 2010 (zie Tabel 1).

7. De toespraken van Verdonk zijn over het algemeen een stuk korter. Dat komt doordat zij een kleinere fractie vertegenwoordigde en dus minder spreektijd kreeg in de debatten.
8. Verwijzingen naar groepen mensen in de samenleving die in oppositie tot 'de kiezer' worden gepresenteerd zijn niet meegerekend (vgl. 'Bijna een derde van alle delicten wordt gepleegd door *niet-westerse allochtonen*.' [W3]); hetzelfde geldt voor verwijzingen naar beroepsgroepen die als voorziening voor kiezers worden opgevoerd (zoals 'thuiszorghulpen' in (11)).
9. De afkortingen achter voorbeelden (zoals [W1] of [V3]) geven aan van wie de voorbeelden afkomstig zijn, en uit welk debat (vgl. Tabel 1). [W1] is een fragment van Wilders uit het Verantwoordingsdebat 2008; [V3] een fragment van Verdonk uit de APB 2008.
10. Een taaluiting is een afgeronde communicatieve eenheid met een illocutionaire strekking die het antwoord geeft op een denkbeeldige vraag (Pander Maat, 2002, pp. 15-22). Er is gekozen voor taaluitingen als analyse-eenheid vanwege de analyse van perspectief (zie de volgende paragrafen): voor een systematische perspectiefanalyse is het noodzakelijk om per communicatieve eenheid vast te stellen of het perspectief van de kiezer wordt opgevoerd, of niet. Een analyse per 'zin' (van hoofdletter tot punt) is daartoe niet toereikend, omdat binnen een zin vaak meerdere taaluitingen kunnen worden onderscheiden die in termen van perspectief van elkaar kunnen verschillen.
11. De keuze om in een geval als (13) de drie verwijzingen als één mee te rekenen (en niet als drie) is ingegeven door de perspectiefanalyse (zie de volgende paragrafen), waarin per taaluiting 'binair' is geanalyseerd of er wel/niet naar het perspectief van kiezers wordt verwezen.
12. Per taaluiting is dus vastgesteld of het perspectief van de kiezer wordt opgevoerd of niet. Een alternatieve analyseprocedure had kunnen zijn om per taaluiting de algemenere vraag te stellen *wiens* perspectief wordt opgevoerd. Zo zou ook inzicht zijn verkregen in de mate waarin Wilders en Verdonk naast zichzelf en kiezers anderen (zoals andere politici) perspectief toekennen. Omdat dit laatste voor de beantwoording van de onderzoeksvraag niet direct relevant is, hebben wij hier niet voor gekozen.
13. Het model van Sanders (1994, p. 67) is iets aangepast om het toe te spitsen op de politieke context. Zo zijn de termen 'verteller' (*narrator*) en 'personage' (*character*) vervangen door 'politicus' en 'kiezer'. Sanders maakt in elke modus (directe, vrije indirecte en indirecte) nog onderscheid tussen rede (*speech*) en gedachte (*thought*). Wij hebben ervoor gekozen om gevallen van perspectivering niet nader uit te splitsen naar 'rede' en 'gedachte' omdat dat voor de doelstellingen van onze studie van ondergeschikt belang was. Tot slot is de laatste stap van Sanders' analysemodel, waarin wordt gevraagd of er nog andere perspectieven in dezelfde uiting te vinden zijn, geschrapt, omdat de analyse er enkel op gericht was om vast te stellen of er in een uiting naar het perspectief van de kiezer wordt verwezen (en zo ja, op welke manier).
14. Voor de statistische berekeningen is gebruik gemaakt van *log likelihood*; de berekeningen zijn uitgevoerd met de log likelihood calculator <<http://ucrel.lancs.ac.uk/llwizard.html>>. Berekening: W: 256/1436 vs. V: 71/686; LL = 18,08, $p < 0,0001$. Zoals uit Tabel 1 blijkt, zijn de speeches van Wilders grofweg dubbel zo lang als die van Vogelaar. Om hiervoor te compenseren, werken we bij de interpretatie van de data met relatieve getallen – zoals in taalkundig-stilistisch onderzoek gangbaar is. Een factor die hiermee niet in de analyse is betrokken, is dat zich in langere speeches wellicht meer *mogelijkheden* voordoen om naar kiezers te verwijzen. Of dat daadwerkelijk zo is, is een punt voor vervolgonderzoek; wij zijn er zelf niet zonder meer van overtuigd, al was het maar omdat wel of niet verwijzen naar kiezers deels een stilistische aangelegenheid is, en daarmee een *keuze* (vgl. Van Leeuwen, 2015, pp. 108-110).

15. Het aantal verwijzingen naar kiezers in Wilders' speeches tijdens de APB van 2008 en 2009 is ook geteld door Van Leeuwen (2015, p. 112). De hier gepresenteerde getallen zijn als gevolg van een andere telmethode (namelijk: per taaluiting – zie de methodesectie) anders, maar ze wijzen in dezelfde richting: uit Tabel 2 blijkt dat Wilders in de APB 2008 meer naar de kiezer verwijst dan in de APB 2009. Dat is in lijn met Van Leeuwens (2015, p. 112) conclusie.
16. Uitzondering hierop vormt van Verdonks speech tijdens het Verantwoordingsdebat 2009.
17. De Algemene Politieke Beschouwingen en de Verantwoordingsdebatten zijn debatten waarvoor relatief veel media-aandacht is en waarmee dus ook in relatief sterke mate een auditorium buiten de Tweede Kamer bereikt kan worden.
18. Loglikelihood-berekening: W: 98/1436 vs. V: 24/686; LL = 9,76, $p < 0,01$.
19. Voorbeelden die Wilders' gebruik van de directe en vrije indirecte rede illustreren, zijn respectievelijk (i) en (ii).
 1. *Een chauffeur uit Gouda* mailde mij vorige week donderdag: "Meneer Wilders, Marokkanen terroriseren ons. Gisteren werd een collega door een Marokkaan tijdens zijn eerste ochtendrit met een mes op zijn keel beroofd, terwijl zijn vriendjes buiten met een auto stonden te wachten. Zij reden er daarna vrolijk vandoor alsof zij naar de bank waren geweest om te pinnen." [W3]
 2. [Dat zijn de onrechtvaardigheden die *mensen in Nederland* mateloos irriteren.] *Je* staat jaren op een wachtlijst en *je* wordt naar achteren geschoven omdat er mensen zoals met een generaal pardon voorgaan. [W1]

Bibliografie

- Akkerman, A., Mudde, C., & Zaslove, A. (2014). How populist are the people? Measuring populist attitudes in voters. *Comparative Political Studies*, 47, 1324-1353.
- Clark, H. (1996). *Using Language*. Cambridge: Cambridge University Press.
- Hawkins, K.A. (2009). Is Chávez populist? Measuring populist discourse in comparative perspective. *Comparative Political Studies*, 42, 1040-1067.
- Jagers, J., & Walgrave, S. (2007). Populism as political communication style: An empirical study of political parties' discourse in Belgium. *European Journal of Political Research*, 46, 319-345.
- Krieken, K. van, Sanders, J., & Hoeken, H. (2015). Viewpoint representation in journalistic crime narratives: An analysis of grammatical roles and referential expressions. *Journal of Pragmatics*, 88, 220-230.
- Leeuwen, M. van (2015). *Stijl en politiek: Een taalkundig-stilistische benadering van Nederlandse parlementaire toespraken*. Utrecht: LOT.
- Moffitt, B. (2016). *The Global Rise of Populism: Performance, Political Style, and Representation*. Stanford, CA: Stanford University Press.
- Oliver, E., & Rahn, W. (2016). Rise of the *Trumpenvolk*: Populism in the 2016 Election. *ANNALS AAPSS*, 667, 189-206.
- Pander Maat, H. (2002). *Tekstanalyse: Wat teksten tot teksten maakt*. Bussum: Coutinho.
- Pauwels, T. (2014). *Populism in Western Europe: Comparing Belgium, Germany and The Netherlands*. London: Routledge.
- Rooduijn, M., & Akkerman, T. (2017). Flank attacks: Populism and left-right radicalism in Western Europe. *Partij Politics*, 23, 193-204.

- Rooduijn, M., de Lange, S., & van der Burg, W. (2014). A populist Zeitgeist? Programmatic contagion by populist parties in western Europe. *Party Politics*, 20, 563-575.
- Rooduijn, M., & Pauwels, T. (2011). Measuring populism: Comparing two methods of content analysis. *West European Politics*, 34, 1272-1283.
- Sanders, J. (1994). *Perspective in Narrative Discourse*. Enschede: Copyprint.
- Sanders, J. (2009). De verdeling van verantwoordelijkheid tussen journalist en nieuwsbron: Vorm en functie van citaatmengvormen in journalistieke genres. *Tijdschrift voor Taalbeheersing*, 31, 1-17.
- Vossen, K. (2009). Hoe populistisch zijn Geert Wilders en Rita Verdonk? Verschillen en overeenkomsten in politieke stijl van twee politici. *Res Publica*, 51, 437-465.

Het effect van monologen en dialogen in radioreclame

Frank van Meurs, Berna Hendriks en Dilek Köksal
Radboud Universiteit Nijmegen

TVT 41 (1): 197–209

DOI: 10.5117/TVT2019.1.013.MEUR

Abstract

The effect of monologues and dialogues in radio commercials

Previous research has shown that the use of dialogues instead of monologues in radio narratives stimulates the imagination of listeners and increases involvement with the narratives (Rodero, 2012). To date, no research has investigated the effectiveness of dialogues versus monologues in radio commercials. The aim of the current study was to investigate the effect of the use of dialogues versus monologues in radio advertisements for different products on involvement with the advertisement, imagery, evaluation of the product, evaluation of the commercial, evaluation of the speaker, and purchase intention. In an experiment with a 2 (presentation method: dialogue, monologue) x 2 (product: shoes, instant coffee) between-subject design, 152 participants evaluated four radio advertisements. Findings showed that dialogues, compared to monologues, led to more lively and attractive radio commercials, a more positive attitude towards the product, a higher purchase intention and higher status of the speaker in the radio commercial. It can be concluded that the use of dialogues instead of monologues can have a positive effect on the effectiveness of radio commercials.

Keywords: dialogue, monologue, radio commercials, persuasion, attitude

Inleiding

Radioreclame is een advertentiemedium dat de adverteerder veel voordelen biedt, zoals relatief lage productiekosten en een groot bereik (Floor &

Van Raaij, 2000, p. 462; Van As, 2000, p. 14). Het grote bereik van radioreclame hangt samen met het feit dat radio een veelbeluisterd medium is, met in Nederland bijvoorbeeld een gemiddelde luistertijd van 2 uur en 39 minuten per dag in de periode oktober-november 2017 (Nationaal Luister Onderzoek, 2017). Radioreclame blijkt gemiddeld drie procent beter beluisterd te worden dan andere programmering, onder andere omdat radioreclame rond het nieuws geprogrammeerd is (RAB, z.j.). Onderzoek heeft laten zien dat een aantal boodschapkenmerken de effectiviteit van radioreclame beïnvloeden (zie voor een overzicht Van As, 2000, pp. 38-46), bijvoorbeeld muziek (Oakes, 2007), het gebruik van mannen- versus vrouwenstemmen (Martín-Santana, Muela-Molina, Reinares-Lara, & Rodríguez-Guerra, 2015) en accent (Hendriks, Van Meurs, & Van der Meij, 2015).

Een mogelijk belangrijk boodschapkenmerk dat nog weinig is onderzocht, is hoe de boodschap in de radioreclame over verschillende sprekers is verdeeld. De boodschap kan worden verteld in de vorm van een monoloog (één spreker) of een dialoog (twee of meer interacterende sprekers: Hmense, 2010, p. 63; Wahl, 2014, p. 246). Beide kunnen gevolgd worden door een voice-over op het einde (Spoelders & Claes, 2006, p. 85), wat Sussex (1989) en Lee (1992) de *comment* na de *action* in de radioreclame noemen. Er zijn verschillende suggesties gedaan over de mogelijke voordelen van het gebruik van dialogen in radioreclame. Spoelders en Claes (2006, p. 75) beargumenteren dat dialogen in radioreclame de aandacht van luisteraars trekken omdat in de dialoog vragen worden opgeworpen waarop zij een antwoord willen hebben. Een website met advies over het schrijven van dialogen voor radioreclame stelt dat een dialoog prikkelend werkt (West, 1999).

Er lijkt weinig experimenteel onderzoek gedaan te zijn naar de effectiviteit van monologen versus dialogen in radioreclame en factoren die daarbij mogelijk een rol spelen. Feingold en Knapp (1977) onderzochten het effect van monologen versus dialogen in radioreclame tegen het gebruik van verdovende middelen. Er bleek in hun studie geen hoofdeffect van dialoog versus monoloog, maar een interactie met het niveau van de dreiging in de boodschap. Wanneer de anti-drugs boodschap een lage dreiging bevatte, werkte de dialoogvorm beter. Wanneer de anti-drugs boodschap een hoge dreiging bevatte, was er nauwelijks verschil tussen de monoloog en de dialoog in de houding van de participanten ten opzichte van verdovende middelen. Het experiment in deze studie had als afhankelijke variabele echter alleen overtuigingskracht (attitude tegenover verdovende middelen), maar onderzocht niet het effect van dialogen versus monologen op andere relevante factoren.

Eerder experimenteel onderzoek naar de effecten van monologen versus dialogen in 'radioverhalen' (audio opnames van een fictief verhaal) liet zien dat dialogen leiden tot meer verbeeldingskracht en meer betrokkenheid bij het verhaal (Rodero, 2012). Hoewel radioverhalen een ander doel hebben dan radioreclames (ontspanning versus overtuiging) en daardoor mogelijk anders verwerkt worden door luisteraars (bijvoorbeeld, met minder resistance to persuasion; Fransen, Smit, & Verlegh, 2015), zouden de effecten die Rodero vond voor het gebruik van dialogen versus monologen in dergelijke verhalen ook relevant kunnen zijn voor radioreclame. Ander onderzoek heeft immers laten zien dat meer verbeeldingskracht in reclame leidt tot verhoogde koopintentie (Adaval & Wyer, 1998; Schlosser, 2003) en dat meer betrokkenheid bij een advertentie leidt tot een betere houding tegenover het product (Polyorat, Alden, & Kim, 2007).

Onderzoek heeft ook laten zien dat, naast betrokkenheid bij de reclame en verbeeldingskracht, de beoordeling van de spreker(s) van belang is bij de evaluatie van radioreclame (Hendriks et al., 2015; Martín-Santana et al., 2015). Het is mogelijk dat de beoordeling van de sprekers verschilt tussen monologen en dialogen. In een monoloog zou een spreker bijvoorbeeld meer gezaghebbend kunnen overkomen omdat hij of zij de hele tijd aan het woord is. In een dialoog zou een spreker bijvoorbeeld sympathieker kunnen overkomen omdat hij of zij op een prettige manier reageert op de andere spreker in de dialoog.

Op grond van bovenstaande overwegingen was onze onderzoeksvraag de volgende: Wat is het effect van dialogen versus monologen in radioreclame op verbeeldingskracht, betrokkenheid bij de reclame, attitude ten opzichte van de reclame, attitude ten opzichte van de spreker, attitude ten opzichte van het product en koopintentie?

Methode

Materiaal

Op grond van Quester en Lin Lim (2003) en Laurent en Kapferer (1985) werd een pretest afgenomen waarin zes producten beoordeeld werden op mate van betrokkenheid: schoenen, horloge, televisie (hoge betrokkenheid), mondwater, ontbijtgranen en oploskoffie (lage betrokkenheid). Betrokkenheid werd gemeten met zes stellingen op zevenpunts Likertschalen (zeer mee oneens – zeer mee eens): 'Ik ben geïnteresseerd in dit product'; 'Dit product is voor mij persoonlijk van belang'; 'Ik ervaar plezier bij het gebruik maken van dit product'; 'Dit product symboliseert

mijn eigenwaarde'; 'Als ik een verkeerde beslissing neem met betrekking tot het product kan dat voor mij belangrijke negatieve consequenties hebben'; 'Ik zou een verkeerde beslissing kunnen nemen met betrekking tot het product' ($\alpha = .89$). Uit de pretest onder 15 proefpersonen bleek dat oploskoffie ($M = 2.71$, $SD = 1.27$) het laagst scoorde op betrokkenheid bij het product en schoenen ($M = 5.52$, $SD = 0.76$) het hoogst.

Voor elk van deze producten werden twee versies van een radioreclame opgenomen. Eén versie werd gescript als monoloog, de andere versie als dialoog. De scripts zijn ontwikkeld en opgenomen in samenwerking met een professioneel radioreclamebureau, RadioHeads. De scripts voor de radioreclames per product waren bijna identiek. Alle radioreclames werden opgenomen door dezelfde professionele stemacteur. In de dialoogconditie nam een tweede mannelijke stemacteur deel aan het gesprek. Het script van de radioreclames is te vinden in Appendix 1. De lengte van de commercials was circa 21 seconden.

Participanten

In totaal beoordeelden 152 proefpersonen de radioreclames (gemiddelde leeftijd: $M = 29.57$; $SD = 11.58$; min 18, max 69; 61.8% vrouw; hoger onderwijs 84.2%). Leeftijd, geslacht en opleidingsniveau waren gelijk verdeeld over de verschillende versies van de radioreclame (leeftijd: $F(3, 148) < 1$; geslacht: $\chi^2(3) = 5.71$, $p = .127$; opleidingsniveau: $\chi^2(18) = 15.08$, $p = .656$). Proefpersonen werden geworven via e-mail en sociale media (Facebook en LinkedIn). Tevens zijn persoonlijke relaties van de derde auteur gevraagd het onderzoek te verspreiden in hun netwerk.

Onderzoeksontwerp

Het onderzoek had een 2 (presentatiewijze: monoloog/dialoog) x 2 (product: oploskoffie/schoenen) tussen-proefpersoonontwerp.

Instrumentatie

Participanten beoordeelden de radiocommercials met behulp van een online vragenlijst op verbeeldingskracht, betrokkenheid bij de advertentie, attitude ten opzichte van het product, attitude ten opzichte van de advertentie, attitude ten opzichte van de spreker en koopintentie. Alle items werden gemeten op 7-punts Likert schalen (1 = zeer mee oneens, 7 = zeer mee eens).

Verbeeldingskracht werd gemeten in termen van duidelijkheid, levendigheid en beeldendheid met zestien items (gebaseerd op Hustinx & De Wit, 2012, de *Imagery Scale* in Rodero, 2012, Ellen & Bone, 1991, en zes

zelfontworpen items) volgend op de stelling 'de radioreclame vond ik...'. Een principale componentenanalyse met oblimin-rotatie liet een oplossing in drie factoren zien die 56.07% van de variantie verklaarden. De drie factoren waren duidelijkheid, levendigheid en beeldendheid. *Duidelijkheid* werd gemeten met de items 'duidelijk, helder, concreet, nuchter, levensecht, dramatisch' ($\alpha = .78$). *Levendigheid* werd gemeten met de items 'levendig, enthousiast, dynamisch, snel gesproken, rustig' ($\alpha = .80$). *Beeldendheid* werd gemeten met de items 'beeldend, gedetailleerd, de radioreclame wekte veel beelden op, ik kon me gemakkelijk iets bij de radioreclame voorstellen, ik zag snel voor me wat er besproken werd in de radioreclame' ($\alpha = .77$).

Betrokkenheid bij de reclame werd gemeten met de *Involvement Scale* van Rodero (2012). Hierbij zijn de volgende vier stellingen gebruikt: 'de radioreclame hield mijn interesse vast', 'de radioreclame hield mijn aandacht vast', 'ik kon me identificeren met het (de) karakter(s) in de radioreclame' en 'de radioreclame wekte bij mij emoties op' ($\alpha = .81$).

Attitude ten opzichte van de reclame werd gemeten in termen van aantrekkelijkheid en begrijpelijkheid met negen items. *Aantrekkelijkheid* werd gemeten met vijf items (gebaseerd op Hendriks et al., 2015; Hornikx, Van Meurs, & Hof, 2013; Lalwani, Lwin, & Li, 2005) volgend op 'ik vind de radioreclame...': 'aantrekkelijk, leuk, origineel, interessant en boeiend' ($\alpha = .91$). *Begrijpelijkheid* werd gemeten met vier items (gebaseerd op Hendriks et al., 2015; Van Meurs, Korzilius, & Hermans, 2004): 'duidelijk, eenvoudig, makkelijk, begrijpelijk' ($\alpha = .77$).

Attitude ten opzichte van de spreker werd gemeten in termen van status en competentie (gebaseerd op Hendriks, Van Meurs, & De Groot, 2017). De inleidende zin was 'de spreker in de radioreclame die het product aanprijst, is/heeft...'. Een principale componentenanalyse met oblimin-rotatie liet een oplossing in twee factoren zien die 56.04% van de variantie verklaarden: status en competentie. *Status* werd gemeten met vijf items: autoriteit, gezaghebbend, dominant, een krachtige stem en zelfverzekerd ($\alpha = .80$). *Competentie* werd gemeten met zes items: betrouwbaar, intelligent, competent, doeltreffend, hoog opgeleid en ambitieus ($\alpha = .83$).

Attitude ten opzichte van het product werd gemeten met drie items (gebaseerd op Hornikx et al., 2013) volgend op 'ik vind het product in de radioreclame...': 'leuk, aantrekkelijk en van goede kwaliteit' ($\alpha = .88$).

Koopintentie werd gemeten met de volgende vier stellingen (gebaseerd op Hendriks et al., 2015; Lalwani et al., 2005): 'dit product is echt iets voor mij'; 'na het beluisteren van de radioreclame...' 'wil ik meer informatie over dit product', 'overweeg ik dit product te kopen' en 'wil ik dit product zeker kopen' ($\alpha = .91$).

Procedure

De vragenlijst werd afgenomen met behulp van een online instrument (Qualtrics). Proefpersonen werden willekeurig verdeeld over de verschillende versies. De gemiddelde duur van het experiment was 6.65 minuten ($SD = 4.05$) De gemiddelde duur per versie verschilde niet ($F(3,148) < 1$).

Resultaten

Gemiddeldes en standaarddeviaties voor alle afhankelijke variabelen worden getoond in Tabel 1.

Tabel 1 Gemiddeldes en standaarddeviaties voor betrokkenheid, verbeeldingskracht, attitudes t.o.v. reclame, spreker, product en koopintentie in functie van presentatiewijze en product (1 = laag, 7 = hoog)

	Dialogoog		Monoloog	
	Schoenen ($n = 39$)	Oploskoffie ($n = 36$)	Schoenen ($n = 39$)	Oploskoffie ($n = 38$)
Verbeeldingskracht				
Duidelijkheid	4.77 (1.14)	4.65 (1.00)	4.65 (1.04)	5.07 (0.86)
Levendigheid	5.11 (0.72)	4.07 (0.87)	3.57 (1.21)	2.98 (0.92)
Beeldendheid	4.65 (0.97)	4.52 (1.19)	4.17 (1.11)	4.63 (1.02)
Betrokkenheid	3.85 (1.28)	3.71 (1.31)	3.19 (1.17)	3.63 (1.37)
Attitude t.o.v. reclame	4.13 (1.37)	3.79 (1.23)	3.38 (1.41)	3.33 (1.37)
Attitude t.o.v. spreker				
Status	4.47 (1.13)	4.49 (1.19)	4.20 (1.17)	3.93 (0.90)
Competentie	4.33 (0.90)	4.51 (1.11)	4.34 (0.96)	4.36 (0.87)
Attitude t.ov. product	4.43 (1.30)	3.70 (1.36)	3.73 (1.36)	3.29 (1.09)
Koopintentie	3.28 (1.44)	2.69 (1.27)	2.71 (1.41)	2.27 (1.08)

Verbeeldingskracht

Uit een tweeweg MANOVA voor duidelijkheid, levendigheid en beeldendheid met als factoren presentatiewijze en product bleek een significant multivariaat effect van presentatiewijze ($F(3, 146) = 25.90, p < .001$) en Product ($F(3, 146) = 11.50, p < .001$). De interactie tussen presentatiewijze en product was niet significant ($F(3, 146) = 1.58, p = .198$).

Uit de univariate analyses bleek alleen een effect van presentatiewijze voor levendigheid ($F(1, 148) = 72.59, p < .001$), maar niet voor duidelijkheid ($F(1, 148) < 1$) of beeldendheid ($F(1, 148) = 1.07, p = .302$). Proefpersonen vonden de dialoog ($M = 4.61, SD = .95$) levendiger dan de monoloog ($M = 3.28, SD = 1.11$).

Uit de univariate analyses bleek een effect van product voor levendigheid ($F(1, 148) = 27.98, p < .001$), maar niet voor duidelijkheid ($F(1, 148) < 1$).

of beeldendheid ($F(1, 148) < 1$). Proefpersonen vonden de radioreclame voor schoenen ($M = 4.34$, $SD = 1.26$) levendiger dan de radioreclame voor oploskoffie ($M = 3.51$, $SD = 1.05$).

Betrokkenheid bij de reclame

Uit een tweeweg ANOVA voor betrokkenheid bij de reclame met als factoren presentatiewijze en product bleken geen significant hoofdeffect van presentatiewijze ($F(1, 148) = 3.24$, $p = .074$) of product ($F(1, 148) < 1$) en geen interactie tussen presentatiewijze en product ($F(1, 148) = 1.96$, $p = .164$).

Attitude ten opzichte van de reclame

Uit een tweeweg ANOVA voor aantrekkelijkheid van de reclame met als factoren presentatiewijze en product bleek een effect van presentatiewijze ($F(1, 148) = 7.73$, $p = .006$), maar geen effect van product ($F(1, 148) < 1$) en geen interactie ($F(1, 148) < 1$). Proefpersonen vonden de dialoog aantrekkelijker ($M = 3.97$, $SD = 1.31$) dan de monoloog ($M = 3.36$, $SD = 1.38$).

Uit een tweeweg ANOVA voor begrijpelijkheid van de reclame bleek geen significant effect van presentatiewijze ($F(1, 148) = 1.75$, $p = .188$), geen effect van product ($F(1, 148) < 1$) en geen interactie ($F(1, 148) = 3.88$, $p = .051$).

Attitude ten opzichte van de spreker

Uit een tweeweg ANOVA voor status met als factoren presentatiewijze en product bleek een effect van presentatiewijze ($F(1, 148) = 5.45$, $p = .021$), maar geen effect van product ($F(1, 148) < 1$) en geen interactie ($F(1, 148) < 1$). Proefpersonen kenden de spreker die het product aanpreeft meer status toe in de dialoog ($M = 4.48$, $SD = 1.15$) dan in de monoloog ($M = 4.07$, $SD = 1.04$).

Uit een tweeweg ANOVA voor competentie met als factoren presentatiewijze en product bleek geen effect van presentatiewijze ($F(1, 148) < 1$), geen effect van product ($F(1, 148) < 1$) en geen interactie ($F(1, 148) < 1$).

Attitude ten opzichte van het product

Uit een tweeweg ANOVA voor attitude ten opzichte van het product bleek een effect van presentatiewijze ($F(1, 148) = 7.15$, $p = .008$) en product ($F(1, 148) = 7.75$, $p = .006$) en geen interactie ($F(1, 148) < 1$). Proefpersonen hadden bij de dialoog een positievere attitude ten opzichte van het product ($M = 4.08$, $SD = 1.37$) dan bij de monoloog ($M = 3.51$, $SD = 1.25$). Proefpersonen hadden een positievere attitude tegenover schoenen ($M = 4.08$, $SD = 1.37$) dan oploskoffie ($M = 3.49$, $SD = 1.24$).

Koopintentie

Uit een tweeweg ANOVA voor koopintentie bleek een effect van presentatiewijze ($F(1, 148) = 5.32, p = .022$) en product ($F(1, 148) = 5.85, p = .017$) en geen interactie ($F(1, 148) < 1$). Proefpersonen hadden bij de dialoog een hogere koopintentie ($M = 2.99, SD = 1.39$) dan bij de monoloog ($M = 2.49, SD = 1.27$). De koopintentie was hoger voor de schoenen ($M = 2.99, SD = 1.45$) dan voor de oploskoffie ($M = 2.47, SD = 1.19$).

Conclusie en discussie

Het doel van dit onderzoek was om na te gaan wat het effect is van dialogen versus monologen in radioreclame op verbeeldingskracht, betrokkenheid bij de reclame, attitude ten opzichte van de reclame, attitude ten opzichte van de spreker, attitude ten opzichte van het product en koopintentie.

De resultaten van ons onderzoek laten zien dat radioreclame met dialoog op de helft van de afhankelijke variabelen beter gewaardeerd werd dan radioreclame met dezelfde informatie, maar dan gepresenteerd als monoloog. Dialogen leidden tot grotere levendigheid van de reclame (een aspect van verbeeldingskracht), een positievere attitude ten opzichte van de reclame, meer status voor de spreker, een positievere attitude ten opzichte van het product en een hogere koopintentie. Er was geen effect van dialoog versus monoloog op twee aspecten van verbeeldingskracht (duidelijkheid en beeldendheid), betrokkenheid bij de reclame, een aspect van attitude ten opzichte van de reclame (begrijpelijkheid) en één aspect van attitude ten opzichte van de spreker (competentie).

Onze bevindingen voor het effect van dialogen versus monologen op verbeeldingskracht komen deels overeen met de bevindingen van Rodero (2012) voor radioverhalen. Hier moet echter wel bij aangetekend worden dat luisteraars mogelijk met een andere verwachting luisteren naar radioverhalen (namelijk als bron van ontspanning) dan naar radioreclame (namelijk als poging van een organisatie om te overtuigen). Onze bevinding dat radioreclame met een dialoog levendiger gevonden wordt dan radioreclame met een monoloog komt overeen met Rodero's bevindingen. Wij vonden echter geen verschillen in duidelijkheid en beeldendheid tussen monologen en dialogen, terwijl Rodero vond dat dialogen wel leiden tot meer duidelijkheid en beeldendheid. Een mogelijke verklaring hiervoor is dat de radioverhalen in Rodero's studie veel langer waren dan onze commercials (vier minuten versus 21 seconden) en dat de dialogen in Rodero's studie tussen zes verschillende personages plaatsvonden en in onze studie

tussen slechts twee personages. Hierdoor was het contrast tussen een lange monoloog en dialoog mogelijk veel prominenter in Rodero's studie, wat tot een groter verschil in duidelijkheid en beeldendheid zou hebben kunnen leiden.

In de huidige studie leidde het gebruik van dialoog niet tot grotere betrokkenheid bij de radioreclame dan het gebruik van monoloog. Dit resultaat is niet in overeenstemming met de bevinding van Rodero (2012) dat een radioverhaal in de vorm van een dialoog, vergeleken met een monoloog, de betrokkenheid bij een radioverhaal verhoogt. Rodero (2012) gaf als verklaring voor de verhoogde betrokkenheid bij het radioverhaal dat de dialoogversie meerdere personages bevatte met wie de luisteraar zich zou kunnen identificeren. Een mogelijke verklaring voor het ontbreken van verhoogde betrokkenheid in de dialoogversie in het huidige onderzoek is dat de dialoogconditie in het onderzoek van Rodero (2012) zes personages bevatte, en die in het huidige onderzoek slechts twee personages, waardoor de mogelijkheid tot identificatie en daarmee betrokkenheid kleiner was.

De bevinding in onze studie dat de attitude tegenover het product beter was in de dialogen dan in de monologen komt deels overeen met de bevinding in Feingold en Knapp (1977) dat de attitude ten opzichte van één van de anti-drugs boodschappen beter was in de dialoog dan in de monoloog. Er was in de studie van Feingold en Knapp geen hoofdeffect van dialoog versus monoloog, maar een interactie effect. Wanneer de anti-drugs boodschap een lage dreiging bevatte werkte de dialoogvorm beter. Daarbij moet worden aangetekend dat het soort overtuigende boodschap in de studie van Feingold en Knapp anders was dan in de huidige studie. In Feingold en Knapp was sprake van een boodschap tegen het gebruik van een product, waarbij het gevaar van het gebruiken van het product werd benadrukt, terwijl in de huidige studie een product werd aangeprezen en gevaar van het gebruik ervan niet werd vermeld. In vervolgonderzoek zou moeten worden nagegaan of het soort overtuigende boodschap een rol speelt bij de effectiviteit van monologen versus dialogen.

De bijdrage van onze studie aan de theorie is dat deze heeft laten zien dat dialogen in een commerciële context – in radioreclame – positieve effecten kunnen hebben op variabelen die direct relevant zijn voor het commerciële doel van de reclame: koopintentie en attitude tegenover het product. Onze studie heeft ook variabelen geïdentificeerd die die grotere overtuigingskracht mogelijk kunnen verklaren: de grotere levendigheid en aantrekkelijkheid van de reclame en hogere status van de spreker in de dialogen.

Naast de variabelen die in deze studie zijn onderzocht, worden er in de literatuur nog andere aspecten genoemd die de grotere overtuigingskracht van dialogen kunnen verklaren. Eén daarvan is dat een dialoog in een radioreclame gezien kan worden als een vorm van *overheard conversation* (Feingold & Knapp, 1977). Bij een monoloog zouden luisteraars het gevoel hebben dat er iemand direct tot hen spreekt om hen te overtuigen, terwijl bij een dialoog de boodschap niet direct tot de luisteraar wordt gericht, maar tot een andere gesprekspartner, die overtuigd wordt. Dat zou ertoe leiden dat bij een dialoog de *resistance to persuasion* (Fransen et al., 2015) bij de luisteraar lager is. Of de weerstand tegen overtuiging bij een dialoog inderdaad lager is dan bij een monoloog, is in het huidige onderzoek – en in eerder onderzoek – niet gemeten. Dat zou in vervolgonderzoek wel moeten gebeuren.

Een ander aspect dat de grotere overtuigingskracht van dialogen zou kunnen verklaren, is dat bij dialogen de luisteraar zich met meerdere sprekers kan identificeren dan bij een monoloog (Lee, 1992; Sussex, 1989). In hoeverre de identificatie met de sprekers bij dialogen inderdaad groter is, is in het huidige onderzoek niet apart gemeten, maar met één item als onderdeel van betrokkenheid bij de reclame. In verder onderzoek zou de invloed van dialogen versus monologen op identificatie met de sprekers expliciet als variabele geoperationaliseerd moeten worden.

Een beperking van het huidige onderzoek is dat bij het inspreken van de radiocommercials alleen gebruik is gemaakt van mannelijke sprekers. Uit eerder onderzoek blijkt dat mannelijke stemmen in radioreclame anders beoordeeld kunnen worden dan vrouwelijke stemmen (Martín-Santana et al., 2015; Rodero, Larrea, & Vázquez, 2012; Whipple & McManamon, 2002). In vervolgonderzoek zouden daarom dialogen en monologen met vrouwelijke sprekers en combinaties van mannelijke en vrouwelijke sprekers onderzocht moeten worden.

Noot

We willen radioreclamebureau RadioHeads hartelijk bedanken voor het maken van de radioreclames die gebruikt zijn in het experiment. Daarnaast gaat onze dank uit naar twee anonieme reviewers voor hun nuttige opmerkingen.

Appendix Scripts Radiocommercials

Reclame Oploskoffie dialoog

SFX (soundeffect): Muziek

Man 1: Man, man, wat ben ik gespannen voor dat sollicitatiegesprek straks.

Man 2: Kijk ik heb de oplossing voor je

gewoon kokend water in dit kopje

een schepje Coffee King erbij

even roeren

word jij weer rustig

Man 1: Mmmm, lekker zeg: even ontspannen met lekkere oploskoffie van Coffee King

Muziek stopt

Reclame Oploskoffie monoloog

SFX (soundeffect): Muziek

Man, man, wat ben ik gespannen voor dat sollicitatiegesprek straks.

Maar ik heb de oplossing

gewoon kokend water in dit kopje

een schepje Coffee King erbij

even roeren

word ik weer rustig

Mmmm, lekker zeg: even ontspannen met lekkere oploskoffie van Coffee King

Muziek stopt

Reclame Schoenen dialoog

SFX (soundeffect): Muziek

Man 1: Moet je eens kijken naar die schoenen! Die kan ik toch niet aan tijdens mijn sollicitatiegesprek?!

Hoe kom ik nou aan schoenen die bij me passen?

Man 2: Kijk! Ik heb de oplossing voor je. Gewoon naar schoendiebijjepast.nl.

Die hebben altijd schoenen die bij je passen

Hier, moet je eens kijken wat een keuze!

Man 1: Ja, perfect! Die andere sollicitanten maken geen schijn van kans!

Muziek stopt

Reclame Schoenen monoloog

SFX (soundeffect): Muziek

Man 1: Moet je eens kijken naar die schoenen! Die kan ik toch niet aan tijdens mijn sollicitatiegesprek?!

Hoe kom ik nou aan schoenen die bij mij passen?

Ik heb de oplossing!

Ik ga gewoon naar schoendiebijjepast.nl. Die hebben altijd schoenen die bij je passen.

Moet je eens kijken wat een keuze!

Perfect! Die andere sollicitanten, ja, die maken geen schijn van kans!

Muziek stopt

Bibliografie

- Adaval, R., & Wyer, R. S. (1998). The role of narratives in consumer information processing. *Journal of Consumer Psychology*, 7(3), 207-245.
- Ellen, P. S., & Bone, P. F. (1991). Measuring communication-evoked imagery processing. In R. H. Holman & M. R. Solomon (Eds.), *Advances in Consumer Research* (pp. 806-812). Provo, UT: Association for Consumer Research.
- Feingold, P. C., & Knapp, M. L. (1977). Anti-drug abuse commercials. *Journal of Communication*, 27(1), 20-28.
- Floor, J. M. G., & Van Raaij, W. F. (2000). *Marketingcommunicatiestrategie*. Leiden: Educatieve Partners Nederland.
- Fransen, M. L., Smit, E. G., & Verlegh, P. W. (2015). Strategies and motives for resistance to persuasion: An integrative framework. *Frontiers in Psychology*, 6, 1201.
- Hendriks, B., Van Meurs, F., & De Groot, E. (2017). The effects of degrees of Dutch accentedness in ELF and in French, German and Spanish. *International Journal of Applied Linguistics*, 27(1), 44-66.
- Hendriks, B., Van Meurs, F., & Van der Meij, E. (2015). Does a foreign accent sell? The effect of foreign accents in radio commercials for congruent and non-congruent products. *Multilingua – Journal of Cross-Cultural and Interlanguage Communication*, 34(1), 119-130.
- Hmensa, P. A. (2010). Radio advertising research: A triangular design perspective from an African context. *Communication Journal of New Zealand*, 11(1), 54-70.
- Hornikx, J., Van Meurs, F., & Hof, R.-J. (2013). The effectiveness of foreign-language display in advertising for congruent versus incongruent products. *Journal of International Consumer Marketing*, 25(3), 152-165.
- Hustinx, L., & De Wit, E. (2012). Kunnen levendige getuigenissen je achter de tralies doen belanden? Een experimenteel onderzoek naar het effect van levendig taalgebruik op oordelen over schuld bij leken en 'experts'. *Tijdschrift voor Taalbeheersing*, 34(3), 213-228.
- Lalwani, A. K., Lwin, M., & Li, K. L. (2005). Consumer responses to English accent variations in advertising. *Journal of Global Marketing*, 18(3/4), 143-165.
- Laurent, G., & Kapferer, J.-N. (1985). Measuring consumer involvement profiles. *Journal of Marketing Research*, 22(1), 41-53.
- Lee, D. (1992). *Competing Discourses: Perspective and Ideology in Language*. London: Routledge.

- Martín-Santana, J. D., Muela-Molina, C., Reinares-Lara, E., & Rodríguez-Guerra, M. (2015). Effectiveness of radio spokesperson's gender, vocal pitch and accent and the use of music in radio advertising. *BRQ Business Research Quarterly*, 18(3), 143-160.
- Nationaal Luister Onderzoek. (2017). Radio luistercijfers oktober-november 2017. Zie http://radio.nl/i/814828/a_image/445/470/oktober-november2017.pdf
- Oakes, S. (2007). Evaluating empirical research into music in advertising: A congruity perspective. *Journal of Advertising Research*, 47(1), 38-50.
- Polyorat, K., Alden, D. L., & Kim, E. S. (2007). Impact of narrative versus factual print ad copy on product evaluation: The mediating role of ad message involvement. *Psychology & Marketing*, 24(6), 539-554.
- Quester, P., & Lin Lim, A. (2003). Product involvement/brand loyalty: Is there a link? *Journal of Product & Brand Management*, 12(1), 22-38.
- RAB. (z.j.). Inzicht in het reclamebereik van radio; reclamebereik is (meer dan) mediumbereik. Zie <http://www.rab.fm/onderzoek/bereiksonderzoek/nlo-reclamebereik/>
- Rodero, E. (2012). Stimulating the imagination in a radio story: The role of presentation structure and the degree of involvement of the listener. *Journal of Radio & Audio Media*, 19(1), 45-60.
- Rodero, E., Larrea, O., & Vázquez, M. (2012). Male and female voices in commercials: Analysis of effectiveness, adequacy for the product, attention and recall. *Sex Roles*, 68(5-6), 349-362.
- Schlosser, A. E. (2003). Experiencing products in the virtual world: The role of goal and imagery in influencing attitudes versus purchase intentions. *Journal of Consumer Research*, 30(2), 184-198.
- Spoelders, S., & Claes, R. (2006). Creative advertisements for the Cinderella medium: The case of Flanders, Belgium. *Journal of Radio Studies*, 13(1), 68-88.
- Sussex, R. (1989). The Americanisation of Australian English: Prestige models in the media In P. Collins & D. Blair (Eds.), *Australian English: The Language of a New Society* (pp. 158-170). Cambridge: Cambridge University Press.
- Van As, J. (2000). *Effectiviteit van radioreclame: Een literatuurstudie naar de effectiviteit van radioreclame en de factoren die deze effectiviteit beïnvloeden*. Amsterdam: SWOCC.
- Van Meurs, F., Korzilius, H., & Hermans, J. (2004). The influence of the use of English in Dutch job advertisements: An experimental study into the effects on text evaluation, on attitudes towards the organization and the job, and on comprehension. *ESP Across Cultures*, 1(1), 93-110.
- Wahl, S. (2014). The contribution of language to multimodal storytelling in commercials. In A. Mariorani & C. Christie (Eds.), *Multimodal Epistemologies: Towards an Integrated Framework* (pp. 245-259). London: Routledge.
- West, B. (1999). Funny you should say that: Five rules for writing radio dialogue. Zie <http://adage.com/article/news/funny-rules-writing-radio-dialogue/92968/>
- Whipple, T. W., & McManamon, M. K. (2002). Implications of using male and female voices in commercials: An exploratory study. *Journal of Advertising*, 31(2), 79-91.

Een significant probleem

Gerben Mulder
Vrije Universiteit Amsterdam

TVT 41 (1): 211–221

DOI: 10.5117/TVT2019.1.014.MULD

Abstract

A significant problem

Null-hypothesis significance testing is the major form of inferential statistics in the field of applied communication research. On many occasions, substantive conclusions are drawn solely on the basis of the (in)significance of a test result. The justification for these conclusions, however, is suspect, because the results of significance testing are too often misinterpreted. The questionable practice of drawing conclusions based on misunderstandings poses a significant problem for the scientific status of communication research. In this contribution the so-called big five misunderstandings (Kline, 2013) of significance testing are described and explained.

Keywords: significance testing, p-values, methodology

Inleiding

Taalbeheersers die kwantitatief onderzoek doen hebben een probleem. Hun bewijsvoering is namelijk nagenoeg volledig gebaseerd op (een vorm van) nul-hypothesetoetsing, terwijl velen de uitkomsten van deze toetsing niet goed interpreteren. Daarin zijn taalbeheersers overigens niet uniek. Ook in aanverwante disciplines zijn misverstanden over statistische toetsing eerder regel dan uitzondering (Haller & Krauss, 2002; Hubbard & Bayarri, 2003; Kline, 2013; Oakes, 1986; Zilliak & McCloskey, 2008).

Wat er in mijn ogen mis is met de bewijsvoering, is dat de conclusies niet gebaseerd worden op de daadwerkelijke resultaten van het onderzoek en logisch

consistente argumentatie, maar vooral op misvattingen over wat de uitkomst van significantietoetsing betekent. Voorbeelden daarvan zijn dat een niet-significant resultaat betekent dat er geen effect is gevonden, dat een alternatieve inhoudelijke hypothese daarom verworpen moet worden, of dat met een significant resultaat de inhoudelijke hypothese is bevestigd. Dit soort conclusies, die dus op grond van misverstanden worden getrokken, zijn onwetenschappelijk, want ongefundeerd. Ongefundeerde conclusies ten aanzien van theorie en praktijk zijn uiteraard ontoelaatbaar voor een wetenschappelijke discipline.

Om een voorbeeld te geven: Mulder (2016) bespreekt hoe misverstanden over de uitkomsten van statistische toetsing hebben geleid tot dubieuze conclusies in zijn proefschrift (Mulder, 2008). Die conclusies kwamen als volgt tot stand. Wanneer er sprake is van een significant experimenteel resultaat (in de juiste richting) dan wordt de conclusie getrokken dat het effect bestaat en wordt de inhoudelijke hypothese aangenomen, maar als er sprake is van een niet-significant toetsresultaat, dan wordt de conclusie getrokken dat er geen effect bestaat. En hoewel Mulder in zijn proefschrift (gelukkig) niet zo ver gaat om als gevolg van niet-significantie zijn inhoudelijke hypothesen expliciet te verwerpen, leidden die niet-significante resultaten wel tot twijfel over de houdbaarheid van de inhoudelijke ideeën. Kortom, de bewijsvoering gebeurt vrijwel exclusief op grond van het al dan niet significant zijn van een toetsresultaat.

De bewijsvoering die Mulder (2016) signaleert is niet uniek. Het is in meer kwantitatief georiënteerd taalbeheersingsonderzoek eerder regel dan uitzondering. Om die reden is de wetenschappelijke status van de conclusies die worden getrokken in dergelijk onderzoek twijfelachtig. Die status hangt immers niet af van de inhoud van de conclusies, maar van de wijze waarop zij tot stand komen.

Het doel van dit artikel is een bijdrage te leveren aan het verbeteren van deze situatie. Ik wil dat doen door misvattingen ten aanzien van significantietoetsing te bespreken en uit te leggen. Ik bespreek wat Kline (2013) de *big five* misverstanden noemt (zie voor een uitgebreide beschrijving van misverstanden en andere cognitieve illusies onder meer Kline, 2013; Goodman, 2008; Lambdin, 2012). Ik zal echter beginnen met het verschaffen van achtergrondinformatie over significantietoetsing.

Een aantal belangrijke toetsingsconcepten

De p-waarde

Bij het uitvoeren van een statistische toets wordt een p-waarde bepaald. Die p-waarde wordt in de praktijk gebruikt om een beslissing te nemen over

de nul-hypothese. Die wordt verworpen als de p-waarde kleiner is dan of gelijk is aan .05 en niet verworpen indien deze groter is dan .05.

De p-waarde wordt bepaald door de waarde van een toetsingsgrootte (zoals t of F) uit te rekenen. De p-waarde is simpelweg de kans om, als de nul-hypothese waar is, een waarde van de toetsingsgrootte aan te treffen die minstens zo groot is als de toetsingsgrootte die is verkregen bij de toetsing. Dus, een resultaat als $t(98) = 2.5$, $p = .0247$, wil zeggen: er is een kans van .0247 om een waarde van $t(98)$ te verkrijgen van 2.5 of groter onder de aanname dat de nul-hypothese waar is.

Een frequentistische interpretatie van het resultaat zou kunnen zijn: bij toetsing van ware nul-hypothesen met een t -toets met 98 vrijheidsgraden wordt in 2.47% van de toetsingen een t -waarde verkregen van 2.5 of groter.

De nul-hypothese

De nul-hypothese is een statistische hypothese. Dat is een theoretische verwachting over alle mogelijke uitkomsten van een onderzoek. Het gaat daarbij om een uitspraak over de kansverdeling van de mogelijke uitkomsten, meestal aangevuld met concrete aannames over parameters in wat informeel een kansverdelingsfunctie kan worden genoemd. De kansverdeling wordt ook wel aangeduid met de term statistische populatie.

Bij de nul-hypothese die bij de t -toets en de F -toets voor het vergelijken van gemiddelden wordt getoetst is de aanname dat alle mogelijke scores met dezelfde kansverdelingsfunctie beschreven kunnen worden, namelijk die van de normale verdeling, ongeacht de groep waarin deze scores worden geobserveerd. Dat houdt in dat de kansverdelingen van alle mogelijke scores in de groepen identiek worden geacht. Vanuit deze statistische nul-hypothese kan de kansverdeling van de toetsingsgrootte deductief worden afgeleid. Dat laatste maakt het mogelijk om de p-waarde uit te rekenen.

Het handige aan een statistische hypothese die veronderstelt dat kansverdelingen identiek zijn is dat daarmee kan worden nagegaan in hoeverre de data consistent zijn met de inhoudelijke hypothese dat verschillen tussen groepen het gevolg zijn van enkel toeval. De vertaling van deze inhoudelijke hypothese is de statistische hypothese dat de kansverdelingen volledig identiek zijn, bijvoorbeeld dat het gaat om identieke normale verdelingen. Mocht er een gegronde reden zijn om te twijfelen aan de aanname van identieke kansverdelingen, dan zou dat kunnen betekenen dat er wél systematische verschillen tussen de groepen zijn.

Maar zelfs als die systematische verschillen er daadwerkelijk zijn, dan impliceert dat natuurlijk helemaal niet dat die verschillen worden veroorzaakt door het fenomeen waarop de inhoudelijke hypothese van de

onderzoeker betrekking heeft. Die inhoudelijke hypothese is één van de mogelijke alternatieve verklaringen. Dat is de reden waarom het verwerpen van de statistische nul-hypothese geen overtuigende evidentie is voor de plausibiliteit van de inhoudelijke hypothese van de onderzoeker. Het verwerpen van een statistische hypothese is dus ook geen afdoende reden om de eigen inhoudelijke hypothese te accepteren.

Type I en Type II fouten

De standaard nul-hypothese is de statistische vertaling van de inhoudelijke aanname dat geobserveerde resultaten (verschillen tussen groepen of correlaties) enkel het gevolg zijn van toeval. De alternatieve hypothese is de statistische vertaling van de aanname dat de observaties niet alleen het gevolg zijn van toeval, maar ook van systematische factoren, bijvoorbeeld de invloed van een experimentele manipulatie. In beide gevallen zijn het statistische hypothesen: theoretische verwachtingen over alle mogelijke resultaten die men kan observeren gegeven de gekozen operationalisering.

Bij (statistische) hypothesetoetsing wordt op basis van de uitkomst van een statistische toets een beslissing genomen over de nul-hypothese. Bij het nemen van die beslissing kunnen twee fouten worden gemaakt. De eerste fout die we kunnen maken is de type I fout. Dat gebeurt als we een ware nul-hypothese verwerpen. In het algemeen vindt men een risico van 5% op een type I fout acceptabel. Die kans moet gezien worden ten opzichte van een oneindig aantal toetsen van ware nul-hypothesen: maximaal 5% van dergelijke toetsingen mag uitmonden in een type I fout. De kans op een type I fout wordt ook wel aangeduid met het symbool α .

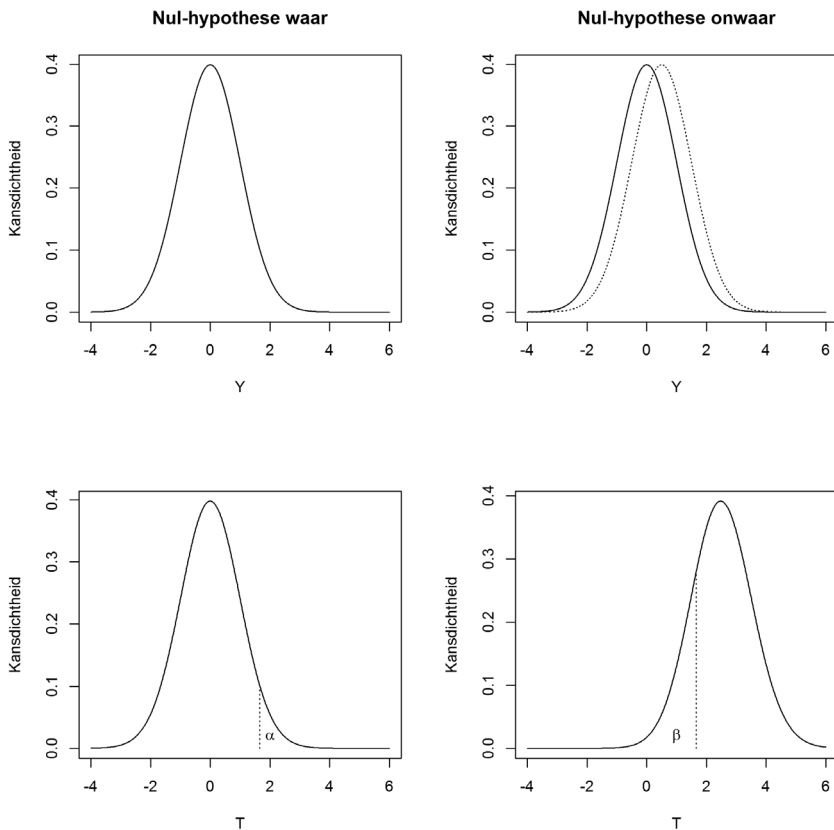
De tweede fout die we kunnen maken is de type II fout, die zich voor doet als we de nul-hypothese niet verwerpen, terwijl de alternatieve hypothese waar is. Om de kans op een type II fout te kunnen bepalen moeten er specifieke waarden worden toegekend aan de parameters van de kansverdelingsfunctie die volgens de alternatieve hypothese geldt. Doorgaans vindt men een risico van 20% acceptabel: maximaal 20% van toetsingen van onware nul-hypothesen mag tot niet-verwerpen leiden.

In Figuur 1 wordt bovenstaande grafisch weergegeven. In de linkerkolom staat de situatie waarin de nul-hypothese waar is. Linksboven zien we een normaal verdeelde populatie van mogelijke scores op de afhankelijke variabele Y . Linksonder staat de kansverdeling van de toetsingsgrootte t . In dit geval geeft de kansverdeling weer wat de verdeling is van de uitkomst van de t -toets wanneer we een oneindig aantal keren een t -toets uitvoeren, telkens op basis van twee nieuwe a -selecte steekproeven uit de normaal verdeelde populatie van de variabele Y . De verticale stippellijn is de waarde

van de kritieke waarde van t . Telkens als we bij de t -toets een waarde van t verkrijgen die groter is dan de kritieke waarde verwerpen we ten onrechte de nul-hypothese. De oppervlakte onder de grafiek rechts van de stippellijn heeft de waarde α , in dit geval 5%. De kans op een type I fout is dus 5%.

In de rechterkolom staat de situatie waarin een specifieke alternatieve statistische hypothese waar is. Die hypothetische situatie in de populatie wordt rechtsboven weergegeven. De hypothese veronderstelt dat er sprake is van twee normale verdelingen die maar op maar één punt verschillen: de parameter μ , het gemiddelde van de verdeling. Dat ligt in de verdeling die is weergegeven met een stippellijn hoger dan bij de verdeling die is weergegeven met een ononderbroken lijn. In dit geval gaat het om wat wel een middelgroot verschil wordt genoemd (Cohen's $\delta = 0.50$).

Rechtsonder wordt de kansverdeling weergegeven van de uitkomst van de t -toets als de alternatieve hypothese waar is. De figuur geeft weer hoe de



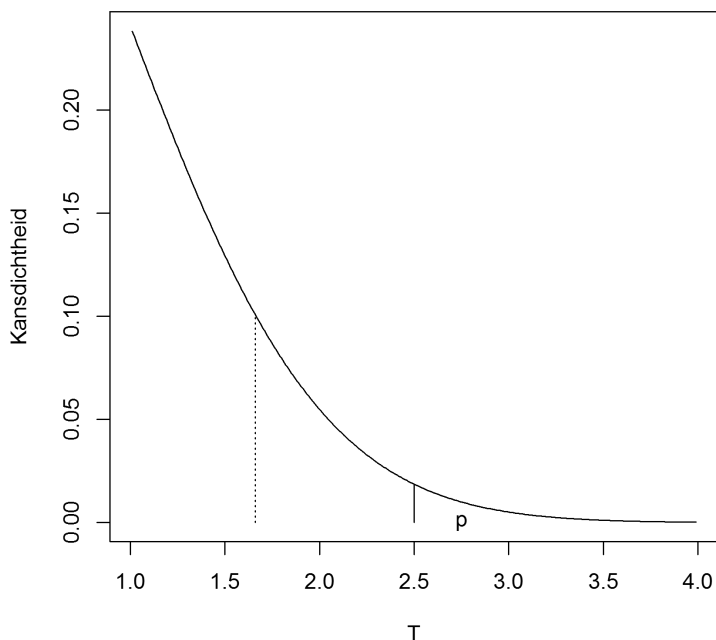
Figuur 1 Kansverdeling van de toetsingsgrootte T bij ware en onware nul-hypothese

verdeling eruit ziet als we een oneindig aantal keren de nul-hypothese toetsen die linksboven wordt weergegeven (bij het toetsen nemen we immers tijdelijk aan dat de nul-hypothese waar is), terwijl we in werkelijkheid te maken hebben met de twee verschillende kansverdelingen die rechtsboven worden weergegeven.

De verticale stippellijn is weer de kritieke waarde van t . Het gebied rechts van die lijn is weer het aantal keren dat we de nul-hypothese verwerpen (in dit geval is dat natuurlijk terecht). Het gebied links van de lijn heeft de waarde β , in dit geval is dat ongeveer 20%. Dat gebied betreft het aantal keren dat we de nul-hypothese ten onrechte niet verwerpen. Oftewel, het betreft de kans dat we een type II fout maken.

Significantie

Van een significant resultaat is doorgaans sprake als de p -waarde van de toets kleiner is dan 5%. In het geval van de t -toets doet die situatie zich voor als we een t -waarde verkrijgen die groter is dan de kritieke waarde van t . In termen van Figuur 1 zou er sprake zijn van een significant resultaat ($p < .05$) als de t -waarde die we krijgen bij de t -toets groter is dan de waarde die correspondeert met de verticale stippellijn. Het volledige gebied rechts van de stippellijn is natuurlijk α .



Figuur 2 Illustratie van een significante p -waarde

In Figuur 2 is ingezoomd op de grafiek linksonder in Figuur 1. Hier is de situatie $t(98) = 2.5$, $p = .0247$ weergegeven. De p-waarde is het gebied rechts van de verticale ononderbroken lijn. Als de nul-hypothese waar is, is er 2.47% kans om bij 98 vrijheidsgraden een t-waarde van 2.5 of groter aan te treffen.

Een significant resultaat wil dus zeggen dat de uitkomst van de statistische toets een waarde heeft die in minder dan 5% van de oneindige hoeveelheid toetsen van ware nul-hypothesen voorkomt.

De *big five* misverstanden

Nu de belangrijke achtergrondinformatie besproken is, is het eenvoudiger om uit te leggen hoe het zit met wat Kline (2013) de *big five* misverstanden noemt.

De odds-against-chance fallacy

De eerste drogreden is dat een significant toetsresultaat het verwerpen van de nul-hypothese rechtvaardigt omdat een significante p-waarde zou betekenen dat de resultaten niet op toeval berusten. De redenering leunt op de misvatting dat de p-waarde op een of andere manier de kans weerspiegelt dat het (toets)resultaat of het effect enkel op toeval berust.

De aanname dat resultaten op toeval berusten is in feite de statistische nul-hypothese. Het idee is dat zelfs als de kansverdelingen identiek zijn, men per toeval verschillen tussen steekproefverdelingen kan aantreffen. Dus als de p-waarde de kans is dat het effect op toeval berust, dan is dat equivalent aan stellen dat de p-waarde de kans is dat de nul-hypothese waar is.

Maar de p-waarde wordt, zoals uitgelegd, berekend op basis van de aanname dat de nul-hypothese waar is. Bij het bepalen wordt dus aangenomen dat de kans 100% is dat de nul-hypothese waar is. Om die reden kunnen we de p-waarde niet zomaar vertalen naar de waarschijnlijkheid dat de nul-hypothese waar is.

De local-type-I-error fallacy

De drogreden is dat verwerpen van een specifieke nul-hypothese gerechtvaardigd is omdat bij verwerpen bij een kleine p-waarde (kleiner dan 5%) de kans klein zou zijn dat je dat ten onrechte doet. Hier wordt ten onrechte gedacht dat wat globaal geldt (t.a.v. de oneindige verzameling toetsen van ware nul-hypothesen) bij verwerpen bij kleine p-waardes ook lokaal geldt: bij één toetsing van één specifieke nul-hypothese. Globaal geldt dat bij

verwerpen bij $p < .05$ en anders niet, een type I risico bestaat van 5%. Maar lokaal geldt dat verwerpen bij $p < .05$ ofwel 100% zeker een type I fout is (de nul-hypothese is waar), ofwel 0% (de nul-hypothese is onwaar), maar op basis van de p-waarde valt niet af te leiden welke van de twee opties het is. Overigens, iemand die gelooft dat verwerpen bij een significant resultaat betekent dat er een kleine kans is op het maken van een type I fout, kan niet anders dan geloven dat een significant resultaat betekent dat er een kleine kans is dat de nul-hypothese waar is. Een type I fout kan immers alleen gemaakt worden als de nul-hypothese waar is.

De inverse-probability-error

De uiteenzetting van de eerste twee misverstanden is al voldoende om te begrijpen waarom de p-waarde niet kan worden gezien als de kans dat de nul-hypothese waar is. Een iets ander perspectief kan desondanks nieuwe inzichten opleveren. De p-waarde is de kans op een waarde van de toetsingsgrootte T die minstens net zo groot is als de verkregen toetsingsgrootte t , bij toetsing van ware nul-hypothesen. Het symbool dat we daarvoor kunnen gebruiken is $p = P(T \geq t \mid H_0)$. Bij de inverse-probability-error wordt die kans omgedraaid, geïnverteerd: $P(H_0 \mid P(T \geq t) = p)$.

In conceptuele termen hebben we, als we de inverse-probability error begaan, te maken met het verwarren van de kans op de data gegeven dat de nul-hypothese waar is met de kans op de nul-hypothese gegeven dat de data waar zijn. Je draait dan twee conditionele kansen om. Een voorbeeld is dat je de kans dat iemand Nederlands verstaat als je weet dat de persoon deelneemt aan het VIOT-congres verwart met de kans dat iemand deelneemt aan het VIOT-congres als deze Nederlands spreekt. De eerste kans is natuurlijk vele malen groter dan de tweede kans.

In het geval van de nul-hypothese weten we precies de kans $p = P(T \geq t \mid H_0)$, maar er is geen mogelijkheid om te bepalen hoe groot de kans $P(H_0 \mid P(T \geq t) = p)$ is. In de theorieën die ten grondslag liggen aan nul-hypothese significantietoetsing is het berekenen van die laatste kans onmogelijk.

De validity-fallacy

De drogreden is dat een significante p-waarde betekent dat de kans groot is dat de (inhoudelijke) alternatieve hypothese waar is en dat het om die reden gerechtvaardigd is de nul-hypothese te verwerpen of de alternatieve hypothese aan te nemen. Het moge duidelijk zijn dat dit een variant is op de bovenstaande drogredenen: als de p-waarde niet vertaald kan worden naar de kans dat de nul-hypothese waar is, kan dat ook niet naar de kans op de alternatieve hypothese.

De replication fallacy

Het misverstand betreft de aanname dat een significant resultaat betekent dat de kans groot is dat het resultaat repliceerbaar is, waarbij repliceren opgevat wordt als dat opnieuw een significant resultaat wordt behaald.

Deze illusie is eenvoudig te weerleggen. Er zijn minstens twee mogelijkheden. Als de nul-hypothese waar is, is de kans dat er een significant resultaat optreedt bij herhaling van het onderzoek gelijk aan 5%. Dus, als er een significant resultaat $p = .05$ wordt behaald terwijl de nul-hypothese waar is, dan is de kans helemaal niet groot dat bij replicatie opnieuw een significant resultaat wordt behaald.

De andere mogelijkheid is dat de nul-hypothese onwaar is. Onder die omstandigheid is de kans dat een significant resultaat wordt behaald de kans $1 - \beta$, de zogeheten *power* van de toets. In het algemeen is die kans onbekend, omdat cruciale informatie, zoals de echte waarde van het effect in de populatie, doorgaans ontbreekt. In elk geval, als de nul-hypothese onwaar is, dan valt op basis van de p-waarde van het eerste onderzoek niet te bepalen hoe groot de kans is dat er een significant resultaat wordt behaald. De gedachte dat bij een significant resultaat de kans op replicatie van het significante resultaat groot is, is dus nergens op gebaseerd.

Conclusie en alternatieven

Een belangrijke consequentie van het bovenstaande is dat een significantietoets niet geschikt is voor het realiseren van de belangrijkste doelen die men met zo'n significantietoets wil bereiken. Die doelen zijn vaststellen of de resultaten op toeval berusten, evalueren hoe waarschijnlijk de inhoudelijke hypothese is en een indicatie verkrijgen van de repliceerbaarheid van resultaten. Dit zijn hele belangrijke vragen, maar een toetsingsprocedure die gebaseerd is op p-waarden levert er (helaas) geen antwoord op.

Maar als significantietoetsing deze vragen niet kan beantwoorden, dan is een vervolgvraag of er (statistische) alternatieven zijn op grond waarvan deze vragen wél beantwoord kunnen worden. Het antwoord is dat er geen magisch statistisch alternatief is waarmee al deze vragen beantwoord kunnen worden (Cohen, 1994), maar dat er wel alternatieven zijn die minder vatbaar zijn voor elementaire misverstanden of die kunnen helpen bij het evalueren van de relatieve plausibiliteit van verschillende statistische hypothesen.

Het enige echte alternatief om te bepalen of resultaten op toeval berusten is replicatie (Kline, 2013). Wanneer je kunt laten zien dat resultaten zich

systematisch voordoen, bijvoorbeeld omdat je een aantal keer een positief effect vindt, dan is dat sterke evidentie dat de resultaten niet op toeval berusten. Zelfs als significantietoetsing bij elk van die replicaties tot een niet-significant toetsresultaat zou leiden.

Een veelgenoemd (frequentistisch) alternatief is te werken met effectgroottes en betrouwbaarheidsintervallen (Cumming & Calin-Jageman, 2017; Mulder, 2016). Zo kun je de effectgrootte en het betrouwbaarheidsinterval gebruiken om te evalueren wat de evidentie zegt over het effect in de populatie. Bijvoorbeeld, als de effectgrootte $d = .05$, en het betrouwbaarheidsinterval is $[-0.05, 0.15]$, dan suggereert dit dat als het effect bestaat, het hoogstens om een zeer klein effect gaat (de limieten van het interval corresponderen met een verwaarloosbaar negatief effect en een zeer klein positief effect).

Een laatste alternatief voor significantietoetsing dat ik wil noemen is de Bayesiaanse benadering van hypothesetoetsing en het schatten van effectgroottes. In tegenstelling tot de technieken die gebaseerd zijn op frequentistische statistische theorie, is het in Bayesiaanse benaderingen mogelijk om de plausibiliteit van statistische hypothesen en specifieke waarden van effectgroottes te kwantificeren (zie Kruschke, 2015 voor een introductie en handleiding), onder meer de kans dat de nul-hypothese waar is.

Bibliografie

- Cohen, J. (1994). The earth is round ($p < .05$). *American Psychologist*, 49, 997-1003.
- Cumming, G. & Calin-Jageman, R. (2017). *Introduction to the New Statistics. Estimation, Open Science, & Beyond*. New York/Londen: Routledge.
- Gigerenzer, G., & Marewski, J. (2015). Surrogate science: The idol of a universal method for scientific inference. *Journal of Management*, 41, 421-440.
- Goodman, S. (2008). A dirty dozen: Twelve p-value misconceptions. *Seminars in Hematology*, 45, 135-140.
- Haller, H., & Krauss, S. (2002). Misinterpretations of Significance. A problem students share with their teachers? *Methods of Psychological Research Online*, 7(1). Zie <https://www.metheval.uni-jena.de/lehre/0405-ws/evaluationuebung/haller.pdf>
- Hubbard, R., & Bayarri, M.J. (2003). Confusion over measures of evidence (p's) versus errors (α 's) in classical statistical testing (with comments). *The American Statistician*, 57, 171-182.
- Kline, R. B. (2013). *Beyond Significance Testing. Statistics Reform in the Behavioral Sciences*. Washington, DC: American Psychological Association.
- Kruschke, J.K. (2015). *Doing Bayesian Data Analysis. A Tutorial with R, Jags, and Stan*. (2nd ed.). London: Academic Press.
- Lambdin, C. (2012). Significance tests as sorcery: Science is empirical – significance tests are not. *Theory & Psychology*, 22, 67-90.
- Mulder, G. (2008). *Understanding Causal Coherence Relations*. Utrecht: LOT. Zie https://www.lot-publications.nl/Documents/172_fulltext.pdf

- Mulder, G. (2016). De kwaliteit van onderzoek. Dichotoom denken versus meta-analytisch denken. *Tijdschrift voor Taalbeheersing*, 38(2), 163-173.
- Oakes, M. (1986). *Statistical Significance*. New York, NY: Wiley.
- Ziliak, S., & McCloskey, D.N. (2008). *The cult of statistical significance: How the standard error costs us jobs, justice, and lives*. Ann Arbor: University of Michigan Press.

De rol van semantische afstand in visuele metaforen

Margot van Mulken, Luuk Lagerwerf en Iris Blokland
Radboud Universiteit Nijmegen/Vrije Universiteit Amsterdam/
PriceWaterhouseCoopers Amsterdam

TVT 41 (1): 223–236

DOI: 10.5117/TVT2019.1.015.MULK

Abstract

The role of semantic distance in visual metaphors

The role of semantic distance between two juxtaposed objects has not been studied thoroughly in visual perception. Does information processing differ between visual metaphors and semantically close objects (hyponyms), or between visual metaphors and semantically distant objects? Probably, semantic distance causes viewers to identify visual metaphors less fast than hyponyms. On the other hand, they might identify visual metaphors faster than objects with similar semantic distance, but without any ground for comparison (ad hoc group). A first experiment with response latencies for identification of 27 object pairs revealed such a pattern, supported by post hoc comprehension measures. In a second experiment, instruction was changed from identification into appreciation. Response latencies were shorter overall. For the ad hoc group, response latencies were longer than for both hyponyms and visual metaphors. Hyponyms were appreciated more than both other groups. Recall was better for visual metaphors than for the ad hoc group. We conclude that a smaller semantic distance indeed helps to identify and appreciate object pairs. Contrary to our expectation, visual metaphors, with a relatively larger semantic distance and a ground for comparison, were not appreciated most.

Keywords: visual metaphor, hyponyms, ad hoc category, semantic distance, juxtaposition, response latencies

Inleiding

Voor de effectiviteit van het genre van advertentie, reclame en voorlichting is het van groot belang om te weten op welke manier kijkers en lezers betekenis geven aan de boodschappen die hen voorgehouden worden. Als we weten hoe de overtuiging tot stand komt, kunnen we met meer succes voorspellen hoe nieuwe boodschappen zullen worden verwerkt. In een reclame-context wordt vaak gewerkt met visuele metaforen – een vorm van beeldspraak waarbij een afbeelding uit een andere context dan die waarvoor geadverteerd wordt overdrachtelijke betekenis heeft (Lagerwerf, Van Hooijdonk, & Korenberg, 2012; Van Mulken, 2003).

Stel, een kijker ziet op een scherm twee objecten naast elkaar, zoals in Figuur 1. Wat gebeurt er als hij¹ besluit dat deze objecten iets met elkaar te maken hebben? De kijker moet besluiten dat de twee objecten niet toevallig naast elkaar staan, en ook dat ze geen loten van dezelfde stam zijn. De kijker moet in feite de semantische afstand tussen de twee objecten bepalen, en beoordelen of het mogelijk is de eigenschappen van het ene object van toepassing te laten zijn op die van het andere object.



Figuur 1 Afbeeldingen van zeep en lijm naast elkaar geplaatst (in juxtapositie)

Het Samenwerkingsbeginsel van Herbert P. Grice gaat niet alleen op voor verbale communicatie maar ook voor visuele communicatie (Grice, 1975). Dat lijkt een triviale constatering, maar bij ons weten is het nog nooit eerder

opgetekend. De deelnemers aan de visuele ‘conversatie’ gaan er beiden van uit dat er impliciete afspraken zijn over de betekenis van de eenheden. Zo gaan ze ervan uit dat de communicatieve bijdrage in Figuur 1 relevante informatie bevat, en niet misleidend is, zodat het de kijker zal lukken de bedoeling van de maker te achterhalen. De twee objecten staan naast elkaar in het verder lege, platte vlak. Ze zijn in juxtapositie, ofwel *aligned* (‘opgelijnd’, Schilperoord, Maes, & Ferdinandusse, 2009). Deze wat kunstmatige context (namelijk de afwezigheid van context, tekst of ander beeld) en de horizontale lijn waarop de objecten naast elkaar zijn geplaatst, zorgen ervoor dat je naar een interpretatie gaat zoeken (Kress & Van Leeuwen, 2006). De oplijning is intentioneel. Het maxime van relevantie en het maxime van stijl maken dat de kijker niet anders kan dan besluiten dat er een bedoeling gesuggereerd wordt, en dat het aan hem is om die bedoeling te achterhalen, omdat hij weet dat de zender erop vertrouwt dat hem dat lukt.

Wat doet de kijker in geval van Figuur 1? Besluit hij dat het gaat om kantoorspullen? Maar hoe past zeep bij kantoorbenodigdheden? Besluit hij dat het gaat om schoonmaakspullen? Maar het eerste object is nu niet direct een schoonmaakmiddel. Misschien besluit hij dat het gaat om een nieuwe categorie: de reparatiecategorie. Repareren wat vuil is en repareren wat kapot is. Als hij besluit tot de reparatiecategorie, dan heeft hij een nieuwe ad hoc categorie gemaakt waarin twee objecten voor het eerst bij elkaar worden gezet (Barsalou, 1983). Of besluit hij dat het gaat om zeepachtige lijm? Om lijm met een geurtje? Om plakkerige zeep? Als hij tot een van de laatste interpretaties besluit, dan heeft de ontvanger een metaforische relatie gelegd: hij heeft een eigenschap van het ene object overgedragen op het andere. Hij heeft dan besloten dat een prototypische karaktertrek van het ene object het andere als het ware besmet door de nabijheid van beide objecten. De meest gebruikte definitie van metaforen is die van Lakoff en Johnson (1980, p. 5): ‘een ding begrijpen in termen van een ander ding.’

In elk geval kan de kijker niet *niet* tot een categorisering proberen te komen – daartoe nodigt de oplijning nu eenmaal uit. Het is wel de vraag of de kijker kan falen in het vinden van een verklaring voor deze oplijning. Sommige auteurs zijn van mening dat een kijker er altijd in slaagt om een voor hemzelf plausibele verklaring te vinden en te interpreteren, vanwege de conceptueel-metaforische structuur van ons denken (Koller, 2009, p. 46; Lakoff & Johnson, 1980). Andere auteurs constateren dat de kijker bij de combinatie van onverenigbare objecten afhaakt (Van Weelden, Maes, Schilperoord, & Cozijn, 2011).

Als een kijker erin slaagt de ‘puzzel op te lossen’, gaat dat gepaard met positief sentiment (Phillips, 1997). De attitude voor de getoonde objecten is daardoor positiever, wat in een reclamecontext kan leiden tot een positieve merk- of productattitude (MacKenzie, Lutz, & Belch, 1986). Ook als kijkers ten onrechte denken dat ze een bevredigende interpretatie van zeeplijm of plakkerige zeep hebben gevonden, zal hun waardering groter zijn dan als ze niets konden bedenken (Lagerwerf, 2002).

In dit artikel onderzoeken we of het gemak waarmee kijkers besluiten dat twee objecten iets met elkaar te maken hebben van invloed is op het begrip, de waardering en herinnering van een afbeelding van twee objecten. Dit doen we om te kunnen vaststellen of visuele metaforen effectiever zijn dan andere juxtaposities – en dat is relevant voor het gebruik van metaforen in een overtuigende context.

Afstand

Veel auteurs die metaforische relaties onderzoeken introduceren een concept dat volgens de één juist groot moet zijn om te kunnen besluiten dat een relatie metaforisch is, en volgens anderen klein. Dat concept heet bij elke auteur weer anders, maar wij scharen het hier onder de noemer ‘semantische afstand’ (vergelijk ook ‘conceptuele gelijkheid’, Van Weelden et al., 2011).

In de Construal Level Theory bepaalt de (psychologische) afstand tussen concepten en een ontvanger het gemak waarmee een metaforische relatie tot stand kan worden gebracht (Jia & Smith, 2013). Voor de hand liggende betekenissen, ofwel *salient meanings* (Giora et al., 2004), geven minder aanleiding om een metafoor te maken dan betekenissen die minder makkelijk uit het geheugen zijn op te halen. Gkiouzepas en Hogg (2011, p.107) stellen dat een grotere ‘conceptuele spanning’ het maken van een metafoor eenvoudiger maakt. Tourangeau en Sternberg (1982) vatten het samen als ‘more distance making better metaphors’.

Andere auteurs beweren juist dat de semantische afstand niet te groot moet zijn, wil een metaforische interpretatie kansrijk zijn. Katz (1989) concludeerde dat bij de beste metaforen het bron- en het doeldomein niet te veel en ook niet te weinig eigenschappen gemeen hadden. Nog verder gaat de conclusie van Johnson en Malgady (1979): hoe meer twee concepten aan elkaar verwant zijn, hoe makkelijker (en beter) de metafoor gevonden werd.

Een complicatie is dat begrip en waardering voor visuele metaforen niet parallel lopen: waardering is het hoogst als begrip niet te makkelijk maar ook niet te moeilijk is (Berlyne, 1970; Van Mulken, Van Hooft, & Nederstigt, 2014). Een vergelijkbare verhouding zou ook kunnen gelden voor semantische afstand: als de semantische afstand toeneemt, neemt begrip af en waardering toe, tot een punt dat de semantische afstand te groot wordt, waardoor begrip blijft afnemen en ook waardering afneemt.

We verwachten dat er een lineair verband is tussen begrip en afstand: als de semantische afstand tussen twee objecten klein is, zal het begrip groter zijn. Voor waardering en herinnering (recall) verwachten we echter een U-curve: de hoogste waardering voor de twee objecten die in semantische afstand niet het verst maar ook niet het kortst van elkaar verwijderd zijn. Hetzelfde verwachten we voor herinnering.

Experiment 1

In een eerste experiment onderzochten we het effect van verschillen in semantische afstand op begrip, waarbij we begrip operationaliseerden als gemak van identificatie (kijktijden) en interpretatie. Als stimuli zijn 27 juxtaposities in drie typen ontwikkeld: hyponiem, metafoor en ad hoc. Drie groepen respondenten kregen elk een andere versie van negen juxtaposities te zien (drie van elk type).

Respondenten

Van de 96 respondenten die begonnen aan het experiment, maakten 65 het af. Hun leeftijd varieerde tussen 16 en 59 jaar ($M = 30,9$ jaar, $SD = 11,9$), en 23,4% was man. Een (V)MBO-opleiding of vergelijkbaar was door 26,6% afgerond, en een HAVO/VWO, HBO- of universitaire opleiding door 73,4%.

Materiaal

Na een pretest bleven 27 paren van afbeeldingen van objecten in juxtapositie over. We maakten hyponiemen (overkoepelende categorie: appel – kers > fruit), metaforen (geen overkoepelende categorie, wel algemeen gedeelde grond voor vergelijking: appel – pil > gezond) en ad hoc-paren (geen overkoepelende categorie, geen algemeen gedeelde grond voor vergelijking: appel – televisie >?; zie Figuur 2). Deze paren werden over drie versies verdeeld, waarbij een object hooguit één keer in een versie voorkwam.



Figuur 2 Voorbeelden van juxtaposities in drie typen: metafoor, hyponiem en ad hoc, met hun overeenkomstsscores uit online WordNet

Semantische afstand tussen de objecten werd berekend met online WordNet semantische-overeenkomstsscores (Postma, Van Miltenburg, Segers, Schoen, & Vossen, 2016). Online WordNet verbindt systematisch betekenseigenschappen van lexicale eenheden. Niet-systematische gronden voor vergelijking zijn geen betekenseigenschappen in WordNet. Metaforische betekenissen tellen dus niet mee in de semantische afstand. We gebruikten WordNet om hyponiemen te onderscheiden: hyponiemen hadden een overeenkomstsscore hoger dan 1,5, en de metaforen en ad hoc-paren lager dan 1,0 (zie Figuur 2).

Instrumentatie

Respondenten werd gevraagd op ‘volgende’ te drukken als ze de afbeelding (met beide objecten) hadden bekeken. Hun kijkduur werd gemeten (Identificatie). Met een open vraag werd de respondenten gevraagd in één tot tien woorden te beschrijven wat de objecten gemeenschappelijk hadden. In de instructie was bovendien duidelijk gemaakt dat ze eventueel ‘geen relatie’ konden invullen. Dit maakte het mogelijk interpretatie te hercoderen in de variabelen Sterke implicatuur en Geen relatie.

Met ‘implicatuur’ bedoelen we een door een kijker afgeleide interpretatie of associatie bij het zien van een combinatie van objecten (Phillips, 1997; Sperber & Wilson, 1986). Er zijn sterke en zwakke implicaturen. Als de meeste respondenten eenzelfde interpretatie geven van een stimulus spreken we van een sterke implicatuur. Er zijn ook zwakke implicaturen: idiosyncratische reacties die tussen individuen juist verschillen (McQuarrie & Phillips, 2005).

Een reactie gold als Sterke implicatuur als 50% of meer van de antwoorden op een stimulus hetzelfde was. Alle andere reacties waren zwakke implicaturen. Daarnaast hebben we de antwoorden geteld waarin werd gemeld dat er geen relatie kon worden benoemd. Dit noemden we de variabele ‘Geen relatie’.

Procedure

Respondenten werden online uitgenodigd deel te nemen aan een Qualtrics-experiment over ‘menselijke creativiteit’. Zij kregen een uitgebreide instructie met drie voorbeelden en voorbeeldvragen. In totaal kregen respondenten negen afbeeldingen te zien (in willekeurige volgorde). Ten slotte vermeldde de respondenten hun leeftijd, geslacht, en opleiding, het apparaat waarmee ze de vragenlijst hadden ingevuld, en of ze waren afgeleid tijdens het experiment. Deelname duurde ongeveer 6 minuten.

Data-analyse

Kijktijden zijn gemeten in milliseconden. Alle kijktijden die hoger uitvielen dan het gemiddelde plus twee keer de standaarddeviatie (per stimulus) zijn verwijderd.

We gebruikten lineaire mixed model analyse met respondenten en stimuli als random factoren en Type Relatie als herhaalde vaste factor. Geslacht en opleiding (in twee groepen, hoog en laag) werden met interactietermen als vaste factor toegevoegd aan het model. Andere factoren leverden geen betere modellen op. De χ^2 -verdeling bepaalde of de informatiecriteria van het lege model (-2RLL waarde) significant verschilden van die van het optimale model, waarbij het verschil in aantal parameters tussen het lege en het optimale model als aantal vrijheidsgraden wordt genomen. Voor Sterke implicaturen is het verschil uitgerekend met alleen hyponiemen en metaforen, en voor Geen relatie met alleen metaforen en ad hoc-paren, omdat er geen sterke implicaturen werden gevonden bij ad hoc-paren, en er geen enkele keer ‘geen relatie’ voorkwam bij hyponiemen. Tabel 1 geeft de informatiecriteria van de drie modellen, met betrekking tot de drie afhankelijke variabelen.

Tabel 1 Informatiecriteria (-2 restricted log likelihood) voor lineaire mixed model analyses van lege modellen (alleen respondenten en stimuli) en Type relatie, opleiding of geslacht toegevoegd als vaste factoren, op Identificatie (kijktijd), Sterke implicaturen en Geen relatie

-2 RLL	Leeg model (parameters)	Optimaal model (parameters)	Factoren
Identificatie	4303,18 (2)	4149,26* (13)	Type relatie, geslacht, opleiding, interacties
Sterke implicatuur	592,79 (2)	579,83* (4)	Type relatie
Geen relatie	444,38 (2)	375,38* (4)	Type relatie
* $p < .01$ (χ^2 -verdeling voor verschil in -2RLL waarden met verschil in parameters als df)			

Resultaten

Tabel 2 geeft de gemiddelden en standaarddeviaties weer van de afhankelijke variabelen voor de drie typen relaties.

Tabel 2 Gemiddelden en standaarddeviaties van Identificatie (kijktijd in s), aantal Sterke Implicaturen, en aantal uitingen van Geen Relatie voor de betekenisrelaties hyponiem, metafoor en ad hoc categorie (N = 65)

	Hyponiem		Metafoor		Ad hoc	
	M	SD	M	SD	M	SD
Identificatie	9,01 ^a	8,02	12,75 ^b	11,27	15,65 ^c	12,54
Sterke implicaturen	0,82 ^a	0,51	0,60 ^b	0,50	-	-
Geen relatie	-	-	0,11 ^a	0,31	0,36 ^b	0,48

Noot: Verschillende superscripten in dezelfde rij duiden op een significant verschil tussen de groepsgemiddelden ($p < 0,05$).

Er was een hoofdeffect van geslacht op Identificatie ($F(1, 530) = 8,24, p = 0,003$); vrouwen ($M = 11,93, SD = 11,05$) keken korter dan mannen ($M = 14,09, SD = 11,15$). Tevens was er een hoofdeffect van opleiding op Identificatie ($F(1, 531) = 18,96, p < 0,001$); de lager opgeleiden ($M = 12,17, SD = 11,54$) keken korter dan de hoger opgeleiden ($M = 12,70, SD = 10,70$). Er waren geen interacties van Type relatie met geslacht ($F < 1$) of opleiding ($F(2, 418) = 1,28, p = 0,28$). Tot slot was er een hoofdeffect van Type relatie ($F(2, 418) = 18,66, p < 0,001$): paarsgewijze vergelijkingen (Bonferroni) toonden aan dat een hyponiem sneller werd geïdentificeerd dan een metafoor ($t(333) = 2,65, p = 0,004$) of een ad hoc categorie ($t(307) = 5,92, p < 0,001$). Daarnaast werd een metafoor sneller herkend dan een ad hoc categorie ($t(357) = 2,61, p = 0,03$).

Er was een hoofdeffect van Type relatie op Sterke Implicaturen: $F(1, 383) = 17,23, p < 0,001$. Een hyponiem heeft meer sterke implicaturen dan een metafoor, maar een ad hoc-paar heeft er helemaal geen (waardoor ad hoc-paren buiten de toetsing vallen).

Daarnaast was er een hoofdeffect van Type relatie op Geen Relatie: $F(1, 331) = 38,93, p < 0,001$. Een metafoor krijgt minder vaak 'geen relatie' als antwoord dan een ad hoc-paar, maar een hyponiem helemaal nooit.

Discussie Experiment 1

In Experiment 1 hebben we onderzocht hoe semantische afstand tussen visuele objecten in juxtapositie het begrip beïnvloedt van respondenten. De verwachting was dat begrip zou afnemen naarmate de afstand groter werd en een grond voor vergelijking ontbrak. Dat lineaire verband werd in kijktijden aangetoond: hyponiemen werden het snelst herkend, metaforen langzamer, en ad hoc-paren het langzaamst. Ook in termen van interpretatie

kwam de verwachting uit: er waren meer sterke implicaturen bij hyponiemen dan bij metaforen, en helemaal geen sterke implicaturen bij de ad hoc-paren. Hetzelfde geldt voor de tweede invulling van interpretatie: niemand ziet 'geen relatie' in de hyponiemen, en minder respondenten zien 'geen relatie' bij metaforen, vergeleken met ad hoc-paren.

Het is mogelijk dat de resultaten het gevolg waren van onze instructies: 'bekijk de plaatjes tot je het begrepen hebt' en 'benoem de relatie tussen de objecten'. Wat zou er gebeuren als we de instructie veranderen? Daarom vragen we in Experiment 2 respondenten naar hun waardering voor (dezelfde) afbeeldingen. Hierbij verwachten we een U-curve, namelijk dat ze de meeste waardering zullen hebben voor metaforen: als kijkers iets meer moeite moeten doen, maar een beloning krijgen in de vorm van het vinden van een vergelijkingsgrond dan zullen ze metaforen meer waarderen dan hyponiemen (want die zijn te makkelijk) of ad hoc-paren (want die zijn te moeilijk).

Experiment 2

In het tweede experiment onderzochten we met hetzelfde materiaal in hetzelfde design en vergelijkbare procedure in hoeverre de respondenten de afbeeldingen waardeerden. Daarbij is ook recall gemeten.

Respondenten

In totaal begonnen 113 respondenten eraan, en 73 maakten het af. Van hen was 61,6% man, en 38,4% vrouw. Zij waren tussen de 17 en 60 jaar oud ($M = 34,45$, $SD = 11,75$). Wat betreft opleiding had 8,2% (V)MBO afgerond, en 91,8% een hogere opleiding (middelbaar, hbo, of universitair).

Instrumentatie

Opnieuw werden kijktijden gemeten. Het verschil met Experiment 1 was dat de respondent wist dat de vervolgtask waardering betrof. In plaats van Identificatie noemen we deze kijktijd daarom Indruk.

Waardering werd gemeten met behulp van een zevenpunts Likertschaal en de deelvragen 'goed', 'aantrekkelijk', en 'plezierig' (Cronbach's $\alpha = 0,85$, MacKenzie et al., 1986).

Na het zien en beoordelen van negen afbeeldingen kregen de respondenten afleidingstaken: zoek-de-verschillen, zet-op-volgorde en een woordspel. Daarna werd gevraagd welke objectcombinaties respondenten zich konden herinneren (Recall). Recall werd gescoord door per respondent en

stimulus te beoordelen of het objectpaar was genoemd (2 punten), een van de twee objecten (1) of geen object (0). Recall werd opgeteld voor elk van de drie condities.

Data-analyse

We gebruikten lineaire mixed-modelanalyse met stimuli en respondent als random factoren, met Type relatie als vaste factor en met Indruk en Waardering als afhankelijke variabelen. Bij de kijktijden voor Indruk zijn tijden langer dan het gemiddelde plus twee maal de standaarddeviatie verwijderd. Omdat de scores voor Recall op geaggregeerd niveau zijn berekend, is daarop een variantieanalyse van herhaalde metingen toegepast.

In Tabel 3 zijn de Informatiecriteria (-2 restricted loglikelihood) weergegeven voor een leeg model (alleen respondenten en stimuli) en voor een optimaal model voor de afhankelijke variabelen Indruk en Waardering. Leeftijd is met een median split in twee groepen verdeeld, andere factoren zijn opleiding (hoger of lager) en geslacht. Geslacht bleek geen relevante factor te zijn in het model, maar leeftijd en opleiding wel, voor Indruk, en opleiding voor Waardering.

Tabel 3 Informatiecriteria (-2 restricted log likelihood) voor lineaire mixed model analyses van lege modellen en met Type relatie op Indruk en Waardering

-2 RLL	Leeg model (parameters)	Optimaal model (parameters)	Factoren
Indruk	3543,39 (2)	3345,80* (12)	Leeftijd, opleiding, Type relatie, interacties
Waardering	2470,32 (2)	2356,48* (9)	Opleiding, Type relatie, interactie

* $p < .01$ (χ^2 -verdeling voor verschil in -2RLL waarden met verschil in parameters als df)

Resultaten

We verwachtten dat de kijktijden voor Indruk hetzelfde patroon zouden volgen als in Experiment 1, en dat Waardering en Recall hoger zouden zijn voor metaforen ten opzichte van de andere twee condities.

In Tabel 4 staan de gemiddelden en standaarddeviaties van Indruk, Waardering en Recall voor de drie typen relaties. Er was een hoofdeffect van leeftijd op Indruk ($F(1, 446) = 3,94, p = 0,048$): de jongere groep ($M = 7,92, SD = 6,71$) keek langer naar de stimuli dan de oudere ($M = 6,28, SD = 3,85$). Een hoofdeffect van opleiding liet zien dat lager opgeleiden ($M = 6,02, SD = 4,24$) korter keken dan hoger opgeleiden ($M = 6,82, SD = 4,07; F(1, 439) = 5,45, p = 0,020$). Er was een marginaal interactie-effect van opleiding en Type relatie ($F(2, 324) = 3,01, p = 0,05$): paarsgewijze vergelijkingen met Bonferronicorrectie toonden aan dat hoger opgeleiden ($M = 6,99, SD = 4,94$)

korter naar de ad hoc-paren keken dan lager opgeleiden ($M = 11,10$, $SD = 10,13$, $t(210) = 2,73$, $p = 0,007$). Voor de andere groepen was er geen verschil ($t < 1$). Verder waren er geen interactie-effecten van leeftijd of opleiding. Er was een hoofdeffect van Type relatie op Indruk ($F(2, 323) = 11,20$, $p < 0,001$). Paarsgewijze vergelijkingen met Bonferronicorrectie lieten zien dat respondenten korter keken naar een hyponiem dan naar een ad hoc-paar ($t(285) = 4,59$, $p < 0,001$), en korter keken naar een metafoor dan naar een ad hoc-paar ($t(351) = 2,84$, $p = 0,01$), terwijl er geen significant verschil in kijktijd was tussen hyponiem en metafoor ($t(362) = 2,16$, $p = 0,09$).

Er was een hoofdeffect van opleiding op Waardering ($F(1, 641) = 4,16$, $p = 0,042$): lager opgeleiden ($M = 4,06$, $SD = 1,82$) gaven een hogere waardering voor de objectparen dan hoger opgeleiden ($M = 3,64$, $SD = 1,56$). Er was geen interactie met Type relatie ($F < 1$). Daarnaast was er een hoofdeffect van Type relatie op Waardering ($F(2, 443) = 25,89$, $p < 0,001$). Paarsgewijze vergelijkingen met Bonferronicorrectie gaven weer dat hyponiemen hoger werden gewaardeerd dan metaforen ($t(424) = 44,43$, $p < 0,001$) en dan ad hoc-paren ($t(432) = 6,75$, $p < 0,001$). Het verschil tussen metaforen en ad hoc-paren was niet significant ($t(431) = 1,19$, $p = 0,71$). De verwachting dat metaforen hoger worden gewaardeerd kwam dus niet uit.

Een variantieanalyse voor herhaalde metingen, met geslacht en Type relatie als factoren, toonde een hoofdeffect van geslacht op Recall ($F(1, 71) = 5,59$, $p = 0,02$, $\eta_p^2 = 0,07$). Mannen ($M = 3,24$, $SD = 1,65$) herinnerden zich minder objectparen dan vrouwen ($M = 3,82$, $SD = 1,52$). Er was geen interactie van geslacht en Type relatie ($F(2, 142) = 1,49$, $p = 0,23$). Ten slotte was er een (klein) hoofdeffect van Type relatie ($F(2, 142) = 3,22$, $p = 0,04$, $\eta_p^2 = 0,04$). Paarsgewijze vergelijkingen met Bonferronicorrectie toonden aan dat metaforen vaker werden genoemd dan ad hoc-paren ($t(70) = 2,54$, $p = 0,04$). Tussen de andere typen relaties was geen verschil (hyponiem-metafoor: $t(70) = 1,39$, $p = 0,50$; hyponiem-ad hoc: $t(70) = 1,14$, $p = 0,77$).

Tabel 4 Gemiddelden en standaarddeviaties van Indruk (kijktijd in s), Waardering (schaal 1-7), en Recall (score 0-6) voor de betekenisrelaties hyponiem, metafoor en ad hoc categorie (N = 73)

	Hyponiem		Metafoor		Ad hoc	
	M	SD	M	SD	M	SD
Indruk	5,32 ^a	2,32	6,53 ^a	3,55	7,32 ^b	5,63
Waardering	4,52 ^a	1,36	3,43 ^b	1,57	3,07 ^b	1,44
Recall	3,41 ^{a,b}	1,55	3,77 ^a	1,64	3,21 ^b	1,69
Noot: Verschillende superscripten in dezelfde rij duiden op een significant verschil tussen de groeps-gemiddelden ($p < 0,05$).						

Discussie en conclusie

Respondenten keken opvallend korter naar objectparen als bleek dat hen om een waardering werd gevraagd dan als ze moest aangeven of ze een relatie tussen beide objecten konden benoemen. Dat alleen al is een opmerkelijke bevinding: veel onderzoek naar het gebruik van (visuele) metaforen in reclames bevraagt enkel de waardering en veronderstelt dan dat waardering ook een resultante van begrip is (naast esthetiek en productevaluatie). Dat zou wel eens veel minder het geval kunnen zijn dan dat we tot nu toe dachten. Wellicht verklaart dat ook waarom we geen evidentie vonden voor onze verwachting dat metaforen in onze tweede studie meer gewaardeerd zouden worden dan hyponiemen of ad hoc-paren (de omgekeerde u-curve): die veronderstelling was gebaseerd op het geïmpliceerde cognitieve plezier van 'moeite doen en oplossing vinden'. De meeste mensen zullen dat pas waarderen als ze de uitdrukkelijke instructie hebben gekregen om moeite te doen. Het is denkbaar dat de principes van de Fluency Theory hier opgeld doen (Reber, Schwarz, & Winkielman, 2004). Het makkelijk kunnen verwerken van een plaatje wordt hoger gewaardeerd dan een complexer plaatje. Onze respondenten keken korter naar hyponiemen, en hadden daar ook een hogere waardering voor, vergeleken met de andere categorieën. Wel bleken ze zich de metaforen beter te kunnen herinneren dan ad hoc-paren. Wellicht dat daar cognitieve inspanning een rol heeft gespeeld, maar het bewijs is dun: er was alleen een klein verschil tussen metaforen en ad hoc.

Veel respondenten haakten af, tijdens het experiment. Nu is dat bij online-experimenten niet ongebruikelijk. Toch is het raadzaam om in de toekomst de deelname aan dit type experimenten beter te surveilleren, bijvoorbeeld door respondenten te belonen of door fysieke aanwezigheid van een laborant.

Is semantische afstand een factor in het interpreteren van visuele metaforen? We onderscheidden hyponiemen van metaforen en ad hoc-paren middels WordNet. We vonden dat een kleinere semantische afstand helpt bij het begrijpen en waarderen van objectparen met een kleinere afstand (hyponiemen). We onderscheidden op basis van vergelijkingsgrond metaforen van de ad hoc-categorie en we vonden inderdaad dat er geen sterke implicaturen gevonden werden bij ad hoc-paren. De verschillen tussen de drie categorieën waren echter niet overal even eenduidig. Niet altijd konden we metaforen van hyponiemen onderscheiden (Indruk experiment 2), niet altijd konden we metaforen van ad hoc-paren onderscheiden (Waardering experiment 2). Bij Recall bleken de metaforen het beter te doen dan de ad

hoc-paren. We hebben dus meer onderzoek nodig. De veronderstelling dat een grotere semantische afstand helpt bij het verwerken van metaforen zou wel eens kunnen kloppen (Gkiouzepas & Hogg, 2011; Jia & Smith, 2013), maar dan moet er wel een herkenbare vergelijkingsgrond zijn, die wordt ingezet voordat kijkers zelf een ad hoc verband gaan verzinnen (Giora et al., 2004). Conventionalisering van vergelijkingsgronden lijkt niet alleen van toepassing te zijn op verbale metaforen (Bowdle & Gentner, 2005), maar ook op visuele.

Wellicht dat een nieuwe studie met hetzelfde materiaal, en een instructie gericht op zowel begrip als waardering (bijvoorbeeld 'kies het beste paar'), meer inzicht brengt. Daarnaast zouden we semantische afstand op andere manieren kunnen operationaliseren. Zodra we weten hoe we een nieuwe, werkzame metafoor kunnen creëren, kunnen we de lees- en kijker-varingen van lezers immers gunstig beïnvloeden.

Noot

1. In navolging van Sperber en Wilson (1986) maakten we de zender vrouwelijk en de ontvanger (kijker) mannelijk.

Bibliografie

- Barsalou, L.W. (1983). Ad hoc categories. *Memory & Cognition*, 11(3), 211-227.
- Berlyne, D.E. (1970). Novelty, complexity and hedonic value. *Perception and Psychophysics*, 8(5a), 279-286.
- Bowdle, B.F., & Gentner, D. (2005). The career of metaphor. *Psychological Review*, 112(1), 193-216.
- Giora, R., Fein, O., Kronrod, A., Elnatan, I., Shuval, N., & Zur, A. (2004). Weapons of mass distraction: Optimal innovation and pleasure readings. *Metaphor and Symbol*, 19(2), 115-141.
- Gkiouzepas, L., & Hogg, M.K. (2011). Articulating a new framework for visual metaphors in advertising. *Journal of Advertising*, 40(1), 103-120.
- Grice, H.P. (1975). Logic and conversation. In P. Cole & J. L. Morgan (Eds.), *Syntax and Semantics. Speech Acts* (Vol. 3, pp. 41-58). New York, NY: Academic Press.
- Jia, L., & Smith, E.R. (2013). Distance makes the metaphor grow stronger: A psychological distance model of metaphor use. *Journal of Experimental Social Psychology*, 49(3), 492-497.
- Johnson, M.G., & Malgady, R.G. (1979). Some cognitive aspects of figurative language: Association and metaphor. *Journal of Psycholinguistic Research*, 8(3), 249-265.
- Katz, A.N. (1989). On choosing the vehicles of metaphors: Referential concreteness, semantic distances, and individual differences. *Journal of Memory and Language*, 28(4), 486-499.
- Koller, V. (2009). Brand images: Multimodal metaphor in corporate branding messages. In C. J. Forceville & E. Urios-Aparisi (Eds.), *Multimodal Metaphor* (Vol. 11, pp. 45-72). Berlin: Mouton de Gruyter.

- Kress, G.R., & Van Leeuwen, T. (2006). *Reading Images: The Grammar of Visual Design*. New York, NY: Routledge.
- Lagerwerf, L. (2002). Deliberate ambiguity in slogans. Recognition and appreciation. *Document Design*, 3(3), 245-260.
- Lagerwerf, L., Hooijdonk, C. M. J., & Korenberg, A. (2012). Processing visual rhetoric in advertisements: Interpretations determined by verbal anchoring and visual structure. *Journal of Pragmatics*, 44(13), 1836-1852.
- Lakoff, G., & Johnson, M. (1980). *Metaphors we Live by*. Chicago, IL: The University of Chicago Press.
- MacKenzie, S.B., Lutz, R.J., & Belch, G.E. (1986). The role of attitude towards the ad as a mediator of advertising effectiveness: A test of competing explanations. *Journal of Marketing Research*, 23, 130-143.
- McQuarrie, E.F., & Phillips, B.J. (2005). Indirect persuasion in advertising. How consumers process metaphors presented in pictures and words. *Journal of Advertising*, 34(2), 7-20.
- Mulken, M. van (2003). Analyzing rhetorical devices in print advertisements. *Document Design*, 4(2), 114-128.
- Mulken, M. van, Van Hooft, A., & Nederstigt, U. (2014). Finding the tipping point: Visual metaphor and conceptual complexity in advertising. *Journal of Advertising*, 43(4), 333-343.
- Phillips, B.J. (1997). Thinking into It: Consumer interpretation of complex advertising images. *Journal of Advertising*, 26(2), 77-88.
- Postma, M., Van Miltenburg, E., Segers, R., Schoen, A., & Vossen, P. (2016). *Open Dutch WordNet*. Paper presented at the Proceedings of the Eighth Global Wordnet Conference, Bucharest, Romania.
- Reber, R., Schwarz, N., & Winkielman, P. (2004). Processing fluency and aesthetic pleasure: Is beauty in the perceiver's processing experience? *Personality and Social Psychology Review*, 8(4), 364-382.
- Schilperoord, J., Maes, A., & Ferdinandusse, H. (2009). Perceptual and conceptual visual Rhetoric: The case of symmetric object alignment. *Metaphor and Symbol*, 24(3), 155-173.
- Sperber, D., & Wilson, D. (1986). *Relevance: Communication and Cognition*. Cambridge, MA: Harvard University Press.
- Tourangeau, R., & Sternberg, R.J. (1982). Understanding and appreciating metaphors. *Cognition*, 11(3), 203-244.
- Weelden, L. van, Maes, A., Schilperoord, J., & Cozijn, R. (2011). The role of shape in comparing objects: How perceptual similarity may affect visual metaphor processing. *Metaphor and Symbol*, 26(4), 272-298.

De rol van gelijkenis bij narratieven

Joëlle Ooms, John Hoeks en Carel Jansen
Rijksuniversiteit Groningen/Universiteit Stellenbosch

TVT 41 (1): 237–250

DOI: 10.5117/TVT2019.1.016.OOMS

Abstract

The role of similarity in narrative persuasion

This paper presents two studies on the role of similarity in the processing of narratives. In Experiment 1 ($N = 83$), female students were presented with one of two versions of a story about the severe consequences of breast cancer. In the similar version, the protagonist was a 22-year-old female student; in the dissimilar version, the protagonist was a 59-year-old working woman. Experiment 2 ($N = 62$) was an online replication of the first study, with an added measure of perceived similarity.

Neither in Experiment 1 nor in Experiment 2 significant differences were found between the two versions on identification, attitude, or intention. Mediation analyses in Experiment 2 did show that the manipulation of similarity had a significant *indirect* influence on identification, transportation, emotions, and intentions, via perceived similarity. These findings suggest that manipulating similarity may influence persuasion, but only when this similarity is perceived as such by the participants.

Keywords: narratives, similarity, transportation, identification, emotions, health communication

Inleiding

Persoonlijke verhalen kunnen overtuigingen, attitudes, intenties en feitelijk gedrag beïnvloeden (Braddock & Dillard, 2016). Daarbij is het van belang dat lezers, c.q. kijkers of luisteraars zich *getransporteerd* voelen in het

verhaal (Green & Brock, 2000) en zich kunnen *identificeren* met de hoofdpersoon (Cohen, 2001). Identificatie en transportatie kunnen mogelijk vergroot worden door de hoofdpersoon van een verhaal te presenteren als gelijkend op de lezer (Slater & Rouner, 2002). Empirisch onderzoek naar de effecten van manipulaties van gelijkenis op transportatie, identificatie, en overtuigingskracht laat echter tegenstrijdige resultaten zien. De Graaf, Sanders en Hoeken (2016) analyseerden elf studies waarin versies van een narratief met een gelijkende of niet-gelijkende protagonist werden vergeleken. In het merendeel van de studies werden objectieve kenmerken gemanipuleerd, zoals geslacht of woonplaats. De Graaf et al. concluderen dat gemanipuleerde gelijkenis in vrijwel geen enkel geval een significante invloed had op transportatie, identificatie of persuasieve uitkomsten zoals attitude en gedragsintentie. Hun conclusie wordt gedeeld door Cohen, Weimann-Saks en Mazor-Tregerman (2017), die geen significante effecten vonden voor hun vier verschillende manipulaties van gelijkenis (nationaliteit en geslacht in studie 1; leeftijd en woonplaats in studie 2). Er zijn echter ook andere, recente studies (bijvoorbeeld Hoeken, Kolthoff, & Sanders, 2016; Igartua, Wojcieszak, Cachón-Ramón, & Guerrero-Martín, 2017) waarin de manipulatie van gelijkenis wel positieve effecten had op identificatie en overtuigingskracht. Het is zelfs zo dat in bijna identieke experimenten verschillende uitkomsten gevonden worden voor gelijkenis. Zo vonden Chen, Bell en Taylor (2016) significant positieve effecten voor een protagonist met dezelfde leeftijd en van hetzelfde geslacht als de deelnemers aan hun onderzoek, terwijl Chen, Bell en Taylor (2017) in een sterk vergelijkbaar experiment geen significante verschillen vonden.

Duidelijk uit dit alles is dat meer onderzoek naar de invloed van gelijkenis van belang is. In twee experimenten onderzochten wij daarom de effecten van gemanipuleerde gelijkenis op identificatie, transportatie, en de persuasieve uitkomsten van narratieven. Ook bepaalden we in beide experimenten de effecten van gemanipuleerde gelijkenis op emoties, die op hun beurt kunnen bijdragen aan narratieve overtuiging. Zo vonden Igartua en Fiuza (2018) dat hun manipulatie van gelijkenis meer negatieve emoties opriep en via die emoties tot een hogere risicoperceptie leidde. Emoties kunnen ook de effecten van transportatie en identificatie op de overtuigingskracht van narratieven mediëren (Green & Brock, 2000).

De eerste onderzoeksvraag was daarom:

- (1) Wat zijn de effecten van gelijkenis op identificatie, transportatie, emoties, en persuasieve uitkomsten?

Om beter zicht te krijgen op het proces dat tot mogelijke effecten van gelijkenis leidt, voegden we in het tweede experiment de volgende onderzoeksvraag toe:

- (2) Medieert waargenomen gelijkenis de relatie tussen gemanipuleerde gelijkenis enerzijds en identificatie, transportatie, emoties, en persuasieve uitkomsten anderzijds?

Experiment 1

Vrouwelijke studenten kregen een van twee versies van een verhaal voorgelegd. In de ene versie leek de protagonist op de studenten, in de andere niet. Na het lezen beantwoordden alle deelnemers vragen over transportatie, identificatie, emoties, attitude en intentie.

Methode

Materiaal. Het materiaal betrof een verhaal over borstkanker, de meest voorkomende vorm van kanker onder vrouwen. De protagonist wordt geconfronteerd met ernstige fysieke en sociale gevolgen van borstkanker en haar toekomst is hierdoor onzeker. In de laatste alinea adviseert de protagonist de lezers om elke maand een zelfonderzoek van de borsten uit te voeren en zo mogelijke kankercellen vroegtijdig op het spoor te komen.

In de twee verhaalversies verschilden de leeftijd van de protagonist en de leefomstandigheden die met leeftijd samenhangen (student versus werkend; vriend versus man; appartement versus huis). In de conditie met de ‘gelijkende’ protagonist werd het verhaal verteld door een 22-jarige vrouwelijke student, in de conditie met de ‘niet-gelijkende’ protagonist ging het verhaal over een 59-jarige werkende vrouw.¹ Zie voor de eerste paragraaf van de twee versies Tabel 1. Beide versies (op te vragen bij de eerste auteur) telden ongeveer 500 woorden.

Tabel 1 Eerste paragraaf van de twee verhaalversies

Gelijkende protagonist	Niet-gelijkende protagonist
Mijn naam is <u>Myrthe</u> . Ik ben <u>22 jaar</u> oud en ik studeer <u>Rechten aan de Rijksuniversiteit Groningen</u> . Tot voor kort dacht ik alles onder controle te hebben. Ik had een fantastische <u>vriend</u> , <u>zicht op</u> een goede baan, een mooi <u>appartement</u> , genoeg geld en veel leuke en trouwe vriendinnen. Wat ik niet onder controle had, was de ontwikkeling van kanker, borstkanker om precies te zijn.	Mijn naam is <u>Gerda</u> . Ik ben <u>59 jaar</u> oud en ik werk <u>parttime in de thuiszorg</u> . Tot voor kort dacht ik alles onder controle te hebben. Ik had een fantastische <u>man</u> , een goede baan, een mooi <u>huis</u> , genoeg geld en veel leuke en trouwe vriendinnen. Wat ik niet onder controle had, was de ontwikkeling van kanker, borstkanker om precies te zijn.

Vragenlijst. Aan alle deelnemers werden dezelfde stellingen voorgelegd, waarmee de volgende variabelen werden bevraagd: transportatie, identificatie, ervaren emoties, en attitude en intentie ten opzichte van het uitvoeren van zelfonderzoek. Alle stellingen werden beoordeeld op een 7-puntsschaal.

Transportatie werd gemeten met zes items. Vijf daarvan waren ontleend aan Moyer-Gusé en Nabi (2010), die zich op hun beurt baseerden op Green en Brock's *Transportation Scale* (2000): 'Ik was mentaal betrokken bij het verhaal terwijl ik het las', 'Tijdens het lezen van het verhaal dwaalden mijn gedachten wel eens af' (omgepoold), 'Het verhaal raakte me', 'Toen ik het verhaal las, vergat ik even alles om me heen', en 'Ik was benieuwd hoe het verhaal zou aflopen'. Daarnaast werd een item van De Graaf, Sanders, Beentjes en Hoeken (2007) gebruikt: 'Mijn aandacht werd volledig opgeslokt door het verhaal' ($\alpha = .87$).

Identificatie met de protagonist werd gemeten met drie van de vijf items die Tal-Or en Cohen (2010) gebruikten: 'Ik begrijp X wel', 'Tijdens het lezen van het verhaal voelde ik de emoties van X', en 'Ik kan me in X verplaatsen' ($\alpha = .77$). De overige twee items waren niet bruikbaar vanwege de relatief geringe lengte van de verhaalversies in deze studie.

Omdat vooral de emoties *angst*, *verdriet*, en *compassie* in eerder onderzoek relevant bleken in de context van persuasieve narratieven over kanker (Ooms, Jansen, Hommes, & Hoeks, 2017) werden deze drie emoties gemeten. Daartoe werd steeds de stelling 'Tijdens het lezen van dit verhaal voelde ik (me) [...] 'gepresenteerd, waarna er voor elke emotie enkele items volgden. Gebaseerd op Dillard, Plotnick, Godbold, Freimuth en Edgar (1996) waren dat: 'angstig', 'bang', 'ongerust' en 'bezorgd' voor angst ($\alpha = .91$), en 'verdrietig', 'bedroefd', en 'somber' voor verdriet ($\alpha = .93$). Gebaseerd op Oliver, Dillard, Bae en Tamul (2012) gebruikten we de items 'medelijden', 'medeleven' en 'sympathie' voor het meten van compassie ($\alpha = .85$).

Attitude ten opzichte van het uitvoeren van borstzelfonderzoek werd gemeten door proefpersonen te vragen hoe 'effectief', 'goed', 'belangrijk', en 'verstandig' ze dit gedrag vonden ($\alpha = .92$). Deze semantische differentiaal werden ontleend aan De Hoog, Stroebe en De Wit (2008), die ook de attitude ten opzichte van een gezondheidsgedraging maten.

Intentie om zelfonderzoek uit te voeren werd gemeten met drie stellingen gebaseerd op Fishbein en Ajzen (2010): 'Ik ben van plan / zal / ga mezelf het komende half jaar maandelijks (te) onderzoeken om borstkanker te ontdekken' ($\alpha = .97$).

Deelnemers en procedure. 83 vrouwelijke studenten van de Rijksuniversiteit Groningen werden willekeurig verdeeld over de condities

‘gelijkende protagonist’ en ‘niet-gelijkende protagonist’. De deelnemers vulden de papieren vragenlijst vrijwillig in tijdens eerstejaarscolleges van verschillende opleidingen aan de Letterenfaculteit.² Hun gemiddelde leeftijd was 20.43 ($SD = 2.24$), lopend van 18 tot 28 jaar.

Resultaten

De gemiddelden en standaarddeviaties voor beide condities staan in Tabel 2. De correlaties tussen de afhankelijke variabelen staan in Tabel 3. Om de effecten van gelijkenis op de gemeten variabelen te bepalen, werd een multivariate variantieanalyse (MANOVA) uitgevoerd. De statistische power ($1-\beta$) van de uitgevoerde toets was, gegeven de omvang van de steekproef ($N = 83$) en $\alpha = .05$, bij een groot effect ($d = .80$) gelijk aan .98, bij een middelgroot effect ($d = .50$) .68, en bij een klein effect ($d = .20$) .11.³

Tabel 2 Gemiddelden (en standaarddeviaties) voor de afhankelijke variabelen, gemeten op een 7-puntsschaal – Experiment 1

	Conditie met gelijkende protagonist ($n = 42$)	Conditie met niet-gelijkende protagonist ($n = 41$)
Identificatie	4.79 (1.17)	4.84 (1.20)
Transportatie	4.51 (1.30)	4.56 (1.24)
Attitude tov zelfonderzoek	5.71 (1.76)	5.98 (1.40)
Intentie tov zelfonderzoek	4.28 (1.79)	3.83 (1.27)
Angst	3.95 (1.25)	3.39 (1.40)
Verdriet	3.72 (1.71)	3.62 (1.61)
Compassie	5.71 (0.98)	5.31 (1.23)

Tabel 3 Correlatiematrix – Experiment 1

	1	2	3	4	5	6	7
1. Identificatie	-						
2. Transportatie	.51**	-					
3. Attitude tov onderzoek	.23*	.16	-				
4. Intentie tov onderzoek	.41**	.39**	.40**	-			
5. Angst	.17	.45**	.11	.25*	-		
6. Verdriet	.27*	.44**	.06	.20	.44**	-	
7. Compassie	.30**	.29**	.23*	.36**	.22	.29*	-

* $p < .05$, ** $p < .01$

Er werd geen significant multivariaat effect gevonden van conditie op de afhankelijke variabelen ($F(7,72) = 1.80$, $p = .10$).

Discussie

In dit experiment lazen vrouwelijke studenten een verhaal met een gelijkende protagonist of een niet-gelijkende protagonist. Er werden geen significante effecten gevonden van de twee verhaalversies. Dit suggereert dat gelijkenis met de protagonist niet van belang is voor identificatie, transportatie, attitude en gedragsintentie.

Opgemerkt moet echter worden dat we in dit experiment niet hebben gemeten in hoeverre onze gelijkenismanipulatie als zodanig is waargenomen. Het kan echter wel nuttig zijn om na te gaan in hoeverre gemanipuleerde boodschapkenmerken worden opgemerkt door deelnemers aan een experiment – niet met het doel een manipulatiecheck uit te voeren (die is strikt genomen onnodig) – maar om mogelijke mediatie-effecten te ontdekken tussen de manipulatie en de afhankelijke variabelen (O’Keefe, 2016; Tao & Bucy, 2007). Een mogelijke beperking van Experiment 1 is verder dat de boodschap op slechts één gedrag was gericht, namelijk het uitvoeren van zelfonderzoek. Onduidelijk is of er wel verschillen in intentie waren gevonden wanneer het aanbevolen gedrag gericht was op het helpen van iemand anders.

Daarom hebben we een nieuw experiment uitgevoerd, waarin metingen van *waargenomen gelijkenis* en de *intentie om te doneren aan een stichting die zich inzet voor kankerbestrijding* zijn toegevoegd.

Experiment 2

Experiment 2 was grotendeels een directe replicatie van Experiment 1. Vrouwelijke studenten kregen hetzelfde materiaal voorgelegd als in Experiment 1. Na het lezen van het verhaal beantwoordden zij vragen over transportatie, identificatie, emoties, attitude en intenties, en nu ook waargenomen gelijkenis.

Methode

Vragenlijst. Naast de variabelen uit Experiment 1 (Cronbach’s α steeds: $> .84$), werden er twee extra variabelen bevraagd: waargenomen gelijkenis en de intentie om te doneren aan een stichting die zich inzet voor kankerbestrijding.

Waargenomen gelijkenis werd gemeten met de *attitude homophily scale* (McCroskey, Richmond, & Daly, 1975): ‘X denkt zoals ik denk’, ‘X gedraagt zich zoals ik me gedraag’, ‘X is hetzelfde als ik’, en ‘X lijkt mij op mij’ ($\alpha = .95$).

Intentie tot doneren werd gemeten met de volgende drie items, opnieuw gebaseerd op Fishbein en Ajzen (2010): 'Ik ben van plan om / ga / zal binnen het komende half jaar (te) doneren aan een stichting die zich inzet voor kankerbestrijding, zoals KWF' ($\alpha = .99$).

Deelnemers en Procedure. Experiment 2 werd online afgenomen met de enquête-tool *Qualtrics*. Via social media werd de link naar de vragenlijst verspreid. In totaal 88 deelnemers, opnieuw vrouwelijke studenten, werden in *Qualtrics* willekeurig toegewezen aan de condities. Omdat het bij online afname niet altijd zeker is dat de materialen grondig genoeg worden gelezen, hebben we onder enkele studenten een pretest uitgevoerd om na te gaan hoe lang deelnemers ten minste nodig zouden hebben om het verhaal te lezen; dat bleek 25 seconden te zijn. Data van deelnemers die korter dan 25 seconden op de pagina met het verhaal waren ($n = 23$) werden daarom niet gebruikt. Dat gold ook voor data van deelnemers die langer dan 180 seconden op de pagina doorbrachten ($n = 3$). De uiteindelijke deelnemersgroep bestond uit 62 vrouwelijke studenten van verschillende opleidingen aan verschillende universiteiten in Nederland (gemiddelde leeftijd = 23.84, $SD = 2.44$, lopend van 18 tot 30 jaar). Met een controlevraag is bevestigd dat de deelnemers verschillend waren dan die aan Experiment 1.

Resultaten

De gemiddelden en standaarddeviaties voor beide condities staan in Tabel 4. De correlaties tussen alle afhankelijke variabelen staan in Tabel 5. De statistische power ($1-\beta$) van de uitgevoerde toets was, gegeven de omvang van de steekproef ($N = 62$) en $\alpha = .05$, bij een groot effect ($d = .80$) gelijk aan .88, bij een middelgroot effect ($d = .50$) .45, en bij een klein effect ($d = .20$) .09.

Tabel 4 Gemiddelden (en standaarddeviaties) voor de afhankelijke variabelen, gemeten op een 7-puntsschaal – Experiment 2

	Conditie met gelijkende protagonist ($n = 33$)	Conditie met niet-gelijkende protagonist ($n = 29$)
Waargenomen gelijkenis	3.52 (1.60)*	2.49 (1.31)*
Identificatie	5.14 (1.24)	4.57 (1.19)
Transportatie	5.01 (1.01)*	4.12 (1.08)*
Attitude tov zelfonderzoek	6.24 (1.15)	6.26 (0.98)
Intentie tov zelfonderzoek	4.12 (1.43)	3.93 (1.48)
Intentie tov doneren	2.78 (1.79)	2.98 (1.65)
Angst	4.92 (1.30)*	4.17 (1.13)*
Verdriet	4.07 (1.82)	4.37 (1.42)
Compassie	5.43 (1.38)	5.32 (1.23)
Noot: Gemiddelden met een asterisk verschillen significant.		

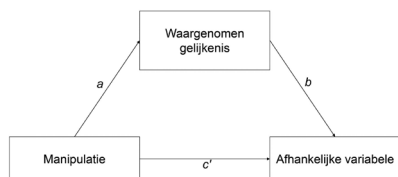
Tabel 5 Correlatiematrix – Experiment 2

	1	2	3	4	5	6	7	8	9
1. Waargenomen gelijkenis	-								
2. Identificatie	.64**	-							
3. Transportatie	.48**	.59**	-						
4. Attitude tov onderzoek	.06	.21	.10	-					
5. Intentie tov onderzoek	.35**	.36**	.09	.31*	-				
6. Intentie tov doneren	.25	.32*	.13	.21	.37**	-			
7. Angst	.33*	.43**	.33**	-.02	.02	-.03	-		
8. Verdriet	.30*	.37**	.35**	-.06	.04	.19	.54**	-	
9. Compassie	.32*	.57**	.36**	.26*	.27*	.30*	.37**	.52*	-

* $p < .05$, ** $p < .01$

Een MANOVA liet een significant multivariaat effect zien van conditie op de afhankelijke variabelen ($F(9,50) = 3.02, p < .01$). Univariate analyses toonden aan dat waargenomen gelijkenis hoger was voor deelnemers die het verhaal lazen over de studente ($M = 3.52$) dan voor deelnemers die het verhaal over de oudere vrouw lazen ($M = 2.49$; $F(1,58) = 7.26, p < .01, \eta^2 = .11$). Er werden geen significante verschillen gevonden tussen de twee condities in identificatie, attitude of intenties. Wel was er een significant verschil in transportatie; deelnemers rapporteerden meer transportatie wanneer zij het verhaal lazen over de studente ($M = 5.01$) dan over de oudere vrouw ($M = 4.12$; $F(1,58) = 9.31, p < .01, \eta^2 = .14$). Van de gemeten emoties bleek alleen angst significant beïnvloed te zijn door conditie ($F(1,58) = 5.98, p < .05, \eta^2 = .09$): deelnemers in de conditie met de gelijkende protagonist ($M = 4.92$) rapporteerden meer angst dan deelnemers in de conditie met de niet-gelijkende protagonist ($M = 4.17$).

Mediatieanalyses werden uitgevoerd met PROCESS (Hayes, 2013). Getoetst werd of waargenomen gelijkenis indirect het effect beïnvloedde van de manipulatie van gelijkenis (0 = ongelijke conditie; 1 = gelijke conditie) op de acht afhankelijke variabelen (zie Figuur 1). Er werden steeds 5000 bootstrap-steekproeven gebruikt.



$a \times b$ indirect effect

c' direct effect

c totaaleffect

Figuur 1 Mediatiemodel

Tabel 6 Uitkomsten mediatieanalyse, met manipulatie van gelijkenis als predictor (X) en waargenomen gelijkenis als mediator (M)

Y:	Pad a	Pad b	Pad c'	Totaal-effect	Indirect effect		
	β	β	β	β	BI 95%		
					β	ondergrens	bovengrens
<i>Transportatie</i>	1.09*	0.27*	0.53	0.82*	0.30	.1206	.6047
<i>Identificatie</i>	1.09*	0.51*	0.05	0.61	0.56	.1869	1.0038
<i>Angst</i>	1.09*	0.22	0.53	0.77*	0.24	-.0610	.7634
<i>Verdriet</i>	1.09*	0.40*	-0.50	-0.06	0.44	.1072	.9811
<i>Compassie</i>	1.09*	0.33*	-0.11	0.24	0.36	.0747	.8306
<i>Attitude tov zelfonderzoek</i>	1.09*	0.04	-0.07	-0.02	0.05	-.0753	.2668
<i>Intentie tov zelfonderzoek</i>	1.09*	0.36*	-0.28	0.12	0.40	.1024	.8869
<i>Intentie tov doneren</i>	1.09*	0.32	-0.48	-0.12	0.35	.0348	.9739

* $p < .01$

Zoals te verwachten gezien de resultaten van de t-toetsen was er slechts in twee van de acht gevallen sprake van een totaaleffect van gemanipuleerde gelijkenis, namelijk op transportatie en angst. Indirecte effecten kunnen echter ook optreden als er geen significant totaaleffect bestaat (Hayes, 2009, pp. 413-414). Zoals te zien in Tabel 6 werden er indirecte effecten gevonden van de manipulatie op transportatie, identificatie, compassie, verdriet, intentie tot zelfonderzoek en intentie tot doneren via waargenomen gelijkenis. De regressiecoëfficiënt van het indirecte effect was steeds positief, net als van de afzonderlijke paden *a* en *b*. Dit betekent dat het verhaal met de gelijke protagonist tot meer waargenomen gelijkenis leidde dan het verhaal met de niet-gelijke protagonist, wat vervolgens leidde tot meer transportatie, identificatie, compassie en verdriet, een sterkere intentie om zelfonderzoek te doen en een sterkere intentie om te doneren.

Discussie

In dit experiment werd onderzocht wat de effecten zijn van gemanipuleerde gelijkenis op identificatie, transportatie, emoties, en persuasieve uitkomsten, en ook op waargenomen gelijkenis. Voor deze laatste variabele werd een significant verschil gevonden: deelnemers vonden zichzelf meer lijken op de studente dan op de oudere vrouw. Net als in Experiment 1 verschilde de mate van identificatie in beide condities echter niet significant. Bij beide verhalen ervoeren de deelnemers een tamelijk hoge mate van identificatie (boven het schaal midden), wat suggereert dat lezers zich

ook kunnen identificeren met een hoofdpersoon waar zij objectief gezien niet op lijken. Terwijl ook de scores voor attitude en beide gemeten intenties niet verschilden voor de twee versies, rapporteerden de deelnemers wel meer transportatie en angst bij het verhaal met de gelijkende protagonist dan bij het verhaal met de niet-gelijkende protagonist.

Experiment 2 werd ook uitgevoerd om de mogelijk mediërende rol van waargenomen gelijkenis te onderzoeken. Hiertoe werden mediatieanalyses uitgevoerd met de manipulatie van gelijkenis als voorspellende variabele, waargenomen gelijkenis als mediërende variabele en de afhankelijke variabelen afzonderlijk als uitkomstvariabelen. Gevonden werd dat waargenomen gelijkenis het effect van de manipulatie van gelijkenis op transportatie, identificatie, compassie, verdriet, intentie tot zelfonderzoek en intentie tot doneren indirect en positief beïnvloedde.

Resultaten Experiment 1 en 2

Om te toetsen of er significante verschillen waren tussen de uitkomsten van de twee experimenten, werd een MANOVA uitgevoerd met conditie en experiment als onafhankelijke variabelen. De statistische power ($1-\beta$) was, gegeven de omvang van de steekproef ($N = 145$) en $\alpha = .05$, bij een groot effect ($d = .80$) gelijk aan 1.00, bij een middelgroot effect ($d = .50$) 1.00, en bij een klein effect ($d = .20$) .32.

Er werden significante multivariate hoofdeffecten gevonden van zowel conditie ($F(7,130) = 2.65, p < .05$) als experiment ($F(7,130) = 3.36, p < .01$). Univariate analyses toonden aan dat het in beide gevallen een significant verschil in angst betrof. Deelnemers in de conditie met de gelijke protagonist ($M = 4.39$) rapporteerden meer angst dan deelnemers in de conditie met de niet-gelijkende protagonist ($M = 3.70$; $F(1,136) = 9.32, p < .01$), en deelnemers in Experiment 2 ervoeren meer angst ($M = 4.57$) dan deelnemers in Experiment 1 ($M = 3.68$; $F(1,136) = 15.24; p < .001$). Er werd geen significant multivariaat interactie-effect gevonden van conditie en experiment ($F(7,130) = 1.78, p = .10$).

Algemene discussie

In dit onderzoek werden de effecten bestudeerd van een manipulatie van objectieve gelijkenis op identificatie, transportatie, emoties, en persuasieve uitkomsten. We hebben twee experimenten uitgevoerd waarin de

protagonist al dan niet leek op de deelnemers wat betreft leeftijd en leefomstandigheden. De experimenten lieten grotendeels dezelfde resultaten zien. Zo werden in geen van beide experimenten effecten gevonden van gemanipuleerde gelijkenis op identificatie. De jonge, vrouwelijke deelnemers ervoeren eenzelfde – tamelijk hoge – mate van identificatie bij de verhaalversie over een protagonist op wie zij leken (i.e., een jonge, studerende vrouw) als bij de versie over een protagonist op wie zij niet leken (i.e., een oudere, werkende vrouw). Blijkbaar is overeenkomst in leeftijd en leefomstandigheden niet nodig om identificatie te ervaren, wat in lijn is met de uitkomsten van Cohen et al. (2017).

Daarnaast leidden de twee onderzochte tekstversies in geen van beide experimenten tot verschillen in attitude of gedragsintentie. In Experiment 2 waren de scores op attitude zodanig hoog dat er mogelijk sprake is geweest van een plafondeffect, maar dat geldt niet voor Experiment 1, en ook niet voor intentie in de twee experimenten. Dat kan dus niet verklaren waarom we geen effecten vinden op attitude en intentie. De Graaf et al. (2016) concluderen in hun systematische review ook dat gelijkenis de overtuigingskracht van narratieven niet beïnvloedt. De onderzoeken die wel positieve effecten vonden voor gelijkenis, bijvoorbeeld die van Hoeken et al. (2016) en Iguarta et al. (2017), manipuleerden gelijkeniskenmerken die relevant waren voor het verhaal. Verhaalrelevante kenmerken hebben wel effect op identificatie en persuasieve uitkomsten in tegenstelling tot kenmerken die niet relevant zijn voor het verhaal (Hoeken et al., 2016; Cohen et al., 2017). Of onze manipulatie van leeftijd verhaalrelevant is, is niet helemaal duidelijk: Een hogere leeftijd vergroot de kans op borstkanker, maar in ons onderzoek werd zowel de oudere als de jongere protagonist getroffen door borstkanker. Misschien waren er wel effecten van gelijkenis gevonden wanneer de manipulatie duidelijker relevant was voor het verhaal, bijvoorbeeld de manipulatie van geslacht bij een verhaal over een kankersoort die alleen bij mannen of alleen bij vrouwen voorkomt.

Anders dan in Experiment 1 vonden we in Experiment 2 ook enkele wel significante verschillen tussen de condities. Bij de versie met de gelijkende protagonist rapporteerden de deelnemers meer transportatie en angst dan bij de versie met de niet-gelijkende protagonist. Mogelijk is het verschil in transportatie te verklaren door de manier van afname. Het kan zo zijn dat deelnemers aan het tweede experiment, dat online werd afgenomen, zich beter konden focussen op het verhaal en het aandachtiger lazen, doordat zij zich bijvoorbeeld thuis bevonden en niet in een collegezaal zoals in het eerste experiment het geval was. In vervolgonderzoek is het dan ook raadzaam om te vragen naar de situatie waarin de deelnemers de vragenlijst

hebben ingevuld, en er zo mogelijk voor te zorgen dat het onderzoek altijd in een rustige omgeving wordt afgenomen.

Hoewel in beide experimenten niet of nauwelijks sprake bleek te zijn van een totaal of direct effect van conditie, bleek de manipulatie van gelijkenis in Experiment 2 wel *indirect* invloed uit te oefenen op transportatie, identificatie, emoties en beide intenties: het verhaal met de gelijkende protagonist leidde tot een hogere waargenomen gelijkenis, wat op zijn beurt een positieve invloed had op transportatie, identificatie, compassie, verdriet, intentie tot zelfonderzoek en intentie tot doneren. Dit suggereert dat positieve effecten van manipulaties van gelijkenis alleen waarneembaar zijn wanneer ontvangers zich ook daadwerkelijk vinden lijken op de protagonist van het verhaal. Vervolgonderzoek naar manieren waarop hoge scores op waargenomen gelijkenis verkregen kunnen worden is dus raadzaam. Mogelijk zorgt specifiekere gelijkenisinformatie tot sterkere effecten van gemanipuleerde gelijkenis, en dan vooral wanneer die informatie de setting van het verhaal betreft, zoals de plaats waar het zich afspeelt (De Graaf et al., 2016, p. 110). Hoewel wij behalve de leeftijd ook de globale leefomstandigheden van de protagonist hebben benoemd (student met appartement versus werkende met huis), moet de setting misschien specifieker beschreven worden om tot grotere verschillen in waargenomen gelijkenis te leiden. Wellicht zou het verhaal ook op andere punten langer gemaakt moeten worden, zodat ontvangers meer tijd hebben om betrokken te raken bij de protagonist (identificatie) en zijn of haar verhaal (transportatie). Er zijn echter geen aanwijzingen in de literatuur dat de duur van het lezen of het bekijken van een narratief de mate van transportatie en identificatie beïnvloedt. Integendeel, narratieve processen, zoals transportatie en identificatie, kunnen ook al in werking treden wanneer ontvangers een afbeelding zien met verhaalelementen (Abbott, 2002). De gemiddelde scores die wij vonden voor transportatie en identificatie zijn boven het middelpunt van de schaal, wat het vermoeden ondersteunt dat een verhaal niet lang hoeft te zijn om betrokkenheid te creëren.

Bij dit alles dient opgemerkt te worden dat wij, net als veel van de eerdere studies naar gelijkenis, slechts één verhaal gebruikt hebben, wat de generaliseerbaarheid van de resultaten niet ten goede komt. Naast vervolgonderzoek met andere verhalen over andere gezondheidsonderwerpen, zou ook een meta-analyse over alle studies – en dat aantal blijft maar toenemen – die gelijkenis hebben gemanipuleerd meer duidelijkheid kunnen verschaffen over de rol van gelijkenis.

In vervolgonderzoek is het verder goed om meer deelnemers per conditie te bevragen, zodat ook eventuele kleine of middelmatige effecten in

de werkelijkheid gevonden kunnen worden. De power van de uitgevoerde analyses in dit onderzoek maakte het alleen goed mogelijk om grote effecten aan het licht te brengen. Met grotere deelnemersaantallen kunnen ook meer geavanceerde pad-analyses uitgevoerd worden, om de onderlinge relaties van de afhankelijke variabelen beter in beeld te brengen en bijvoorbeeld te toetsen of transportatie via emoties de overtuigingskracht beïnvloedt (Green & Brock, 2000). Zulke analyses hebben we inmiddels wel kunnen uitvoeren in een vervolgstudie naar de effecten van gemanipuleerde gelijkenis. Daarover hopen we binnenkort verslag te doen.

Noten

1. In beide experimenten werd ook een tekstversie gebruikt over de mogelijke gevolgen van borstkanker op vrouwen in het algemeen, zonder dat leeftijd of andere kenmerken van de protagonist werden genoemd. In deze tekst kwam geen specifieke hoofdpersoon aan het woord, waarmee niet werd voldaan aan de definitie van een narratief als 'a representation of events involving one or more *characters*' (zie bijvoorbeeld Abbott, 2002). Aangezien in dit artikel de rol van gelijkenis in *narratieven* centraal staat, vergelijken we hier alleen de tekstversies met een gelijkende of niet-gelijkende protagonist.
2. Met dank aan masterstudent Sebastiaan Hartman voor de dataverzameling in Experiment 1.
3. De poweranalyses werden uitgevoerd met behulp van het programma G*Power 3.1 (Faul, Erdfelder, Buchner, & Lang, 2007). Voor de grenswaarden van *d*, zie Cohen (1977).

Bibliografie

- Abbott, H.P. (2002). *The Cambridge Introduction to Narrative*. Cambridge: University Press.
- Braddock, K., & Dillard, J.P. (2016). Meta-analytic evidence for the persuasive effect of narratives on beliefs, attitudes, intentions, and behaviors. *Communication Monographs*, 83(4), 446-467.
- Chen, M., Bell, R.A., & Taylor, L.D. (2016). Narrator point of view and persuasion in health narratives: The role of protagonist-reader similarity, identification, and self-referencing. *Journal of Health Communication*, 21(8), 908-918.
- Chen, M., Bell, R.A., & Taylor, L.D. (2017). Persuasive effects of point of view, protagonist competence, and similarity in a health narrative about type 2 diabetes. *Journal of Health Communication*, 22(8), 702-712.
- Cohen, J. (1977). *Statistical power analysis for the behavioral sciences* (Rev. ed.). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.
- Cohen, J. (2001). Defining identification: A theoretical look at the identification of audiences with media characters. *Mass Communication & Society*, 4(3), 245-264.
- Cohen, J., Weimann-Saks, D., & Mazor-Tregerman, M. (2017). Does character similarity increase identification and persuasion? *Media Psychology*, 21(3), 506-528.
- De Graaf, A., Sanders, J., Beentjes, H., & Hoeken, H. (2007). De rol van identificatie in narrative overtuiging. *Tijdschrift voor Taalbeheersing*, 29(3), 237-250.

- De Graaf, A., Sanders, J., & Hoeken, H. (2016). Characteristics of narrative interventions and health effects: A review of the content, form, and context of narratives in health-related narrative persuasion research. *Review of Communication Research*, 4, 88-131.
- De Hoog, N., Stroebe, W., & De Wit, J.B.F. (2008). The processing of fear-arousing communications: How biased processing leads to persuasion. *Social Influence*, 3(2), 84-113.
- Dillard, J.P., Plotnick, C.A., Godbold, L.C., Freimuth, V.S., & Edgar, T. (1996). The multiple affective outcomes of AIDS PSAs: Fear appeals do more than scare people. *Communication Research*, 23(1), 44-72.
- Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, 39, 175-191.
- Fishbein, M., & Ajzen, I. (2010). *Predicting and Changing Behavior: The Reasoned Action Approach*. New York, NY: Psychology Press.
- Green, M.C., & Brock, T.C. (2000). The role of transportation in the persuasiveness of public narratives. *Journal of Personality and Social Psychology*, 79(5), 701-721.
- Hayes, A.F. (2009). Beyond Baron and Kenny: Statistical mediation analysis in the new millennium. *Communication Monographs*, 76(4), 408-420.
- Hayes, A.F. (2013). *Introduction to Mediation, Moderation, and Conditional Process Analysis: A Regression-based Approach*. New York, NY: Guilford Publications.
- Hoeken, H., Kolthoff, M., & Sanders, J. (2016). Story perspective and character similarity as drivers of identification and narrative persuasion. *Human Communication Research*, 42(2), 292-311.
- Igartua, J.J., Wojcieszak, M., Cachón-Ramón, D., & Guerrero-Martín, I. (2017). "If it hooks you, share it on social networks". Joint effects of character similarity and imagined contact on the intention to share a short narrative in favor of immigration. *Revista Latina de Comunicación Social*, 72, 1085-1106.
- Igartua, J.J., & Fiuza, D. (2018). Persuading with narratives against gender violence. Effect of similarity with the protagonist on identification and risk-perception. *Palabra Clave*, 21(2), 499-523.
- McCroskey, J.C., Richmond, V.P., & Daly, J.A. (1975). The development of a measure of perceived homophily in interpersonal communication. *Human Communication Research*, 1(4), 323-332.
- Moyer-Gusé, E., & Nabi, R.L. (2010). Explaining the effects of narrative in an entertainment television program: Overcoming resistance to persuasion. *Human Communication Research*, 36, 26-52.
- O'Keefe, D.J. (2016). *Persuasion: Theory and Research* (3rd ed.) Los Angeles, CA: Sage.
- Oliver, M.B., Dillard, J.P., Bae, K., & Tamul, D.J. (2012). The effect of narrative news format on empathy for stigmatized groups. *Journalism & Mass Communication Quarterly*, 89(2), 205-224.
- Ooms, J.A., Jansen, C.J.M., Hommes, S., & Hoeks, J.C.J. (2017). "Don't Make My Mistake": On the Processing of Narrative Fear Appeals. *International Journal of Communication*, 11, 4924-4945.
- Slater, M.D. & Rouner, D. (2002). Entertainment-education and elaboration likelihood: Understanding the processing of narrative persuasion. *Communication Theory*, 12(2), 173-191.
- Tal-Or, N., & Cohen, J. (2010). Understanding audience involvement: Conceptualizing and manipulating identification and transportation. *Poetics*, 38, 402-418.
- Tao, C.-C., & Bucy, E.P. (2007). Conceptualizing media stimuli in experimental research: Psychological versus attribute-based definitions. *Human Communication Research*, 33(4), 397-426.

Retorische vragen van de rechter als wrakingsgrond

Een argumentatie-analytisch model

H. José Plug

Universiteit van Amsterdam

TVT 41 (1): 251–264

DOI: 10.5117/TVT2019.1.017,JOSE

Abstract

The judge's rhetorical questions as ground of objection: an argumentation-analytical model

A complaint about the lack of impartiality of a judge may be filed on the grounds of the judge's verbal behaviour. Such a complaint may result in the disqualification of the judge. In this paper, criticism on the verbal behaviour of the judge will be discussed from an argumentative perspective. From this perspective, the disqualification procedure may be considered a sub-discussion that relates to the initial legal dispute. I will examine what role complaints about the verbal behaviour of the judge can play in the argumentation of a party who seeks the judge's disqualification. Therefore, I will analyse the structuring of the argumentation underlying disqualification requests that are based on complaints about the verbal behaviour of the judge. After that I will concentrate on complaints that concern the judge's use of rhetorical questions. How these complaints about rhetorical questions may be evaluated will be examined and illustrated by an argumentative analysis of a decision by a disqualification panel.

Keywords: derailment of strategic manoeuvring, impartiality, judicial disqualification, legal argumentation, strategic manoeuvring, rhetorical question

Inleiding

In het recht is bepaald dat een procespartij de rechter kan wraken als er bij die partij twijfel bestaat over de onpartijdigheid van de betrokken rechter. Klachten over het vermeende gebrek aan onpartijdigheid van de rechter kunnen betrekking hebben op zowel het non-verbale gedrag als het verbale gedrag van de rechter. Uit onderzoek van onder andere Van Rossum, Tigchelaar en Ippel (2012, pp. 24-29) blijkt dat het aantal gevallen waarin rechters worden gewraakt is toegenomen. Deze toename wordt mede verklaard uit de veranderde (communicatieve) rol van de rechter. Traditioneel is het de taak van de rechter om de wet toe te passen waarbij zijn of haar rol zich beperkt tot die van 'spreekbuis van de wet'. Echter, de recente heroriëntatie op de rol van de rechter vereist een actiever en communicatiever optreden tijdens rechtszittingen (Bauw, 2011; Plug, 2016a). Dit betekent onder andere dat de rechter zich niet langer uitsluitend concentreert op de juridische analyse van een zaak maar met de procespartijen in dialoog gaat over de meest adequate manier om de zaak tot een goed einde te brengen. De actievere en communicatievere rol van de rechter kan weliswaar een positief effect hebben op de uitkomst van het geschil en de manier waarop de procespartijen de juridische procedure ervaren, maar het verhoogt het risico voor de rechter om te worden bekritiseerd op zijn verbale gedrag. Wanneer die kritiek resulteert in een wrakingsverzoek, beslist de wrakingskamer of het vermeende gebrek aan onpartijdigheid kan worden afgeleid uit het verbale gedrag van de rechter gedurende de rechtszitting.

In deze bijdrage onderzoek ik hoe kritiek op het verbale gedrag van de rechter, in het bijzonder op het gebruik van retorische vragen, een rol kan spelen in het strategisch manoeuvreren van de procespartij die verzoekt om wraking van de rechter. Om te kunnen vaststellen of en wanneer het strategisch manoeuvreren ontspoot, introduceer ik een model dat kan dienen als vertrekpunt voor de analyse en evaluatie van strategische manoeuvres met betrekking tot retorische vragen die de onpartijdigheid van de rechter betwisten.

Voor het onderzoek naar het strategisch manoeuvreren in wrakingsprocedures maak ik gebruik van de uitgebreide pragma-dialectische argumentatietheorie (Van Eemeren, 2010). Vanuit dit theoretisch perspectief bezien, worden de argumentatieve bijdragen van de rechter en die van de procespartijen opgevat als manifestaties van strategisch manoeuvreren. Omdat zowel de procespartijen als de rechter worden beschouwd als deelnemers aan een rationele, oplossingsgerichte discussie, mag ervan worden uitgegaan dat zij strategisch zullen manoeuvreren binnen de grenzen van

de redelijkheid. Dit betekent dat zij in principe hun dialectische en retorische doelen met elkaar in evenwicht brengen en de regels voor een redelijke discussie respecteren.

In de volgende paragraaf geef ik een argumentatieve reconstructie van een wrakingsprocedure als een sub-discussie die uitmondt in het besluit een wrakingsverzoek al dan niet toe te wijzen. Daarna laat ik zien hoe het gebruik van een retorische vraag door de rechter kan resulteren in een klacht over een (vermeend) gebrek aan onpartijdigheid en, als zodanig, in een beschuldiging van ontspoord strategisch manoeuvreren door de rechter. Op basis van deze analyse stel ik een aantal theoretisch onderbouwde criteria voor die als uitgangspunt kunnen dienen voor de analyse en evaluatie van dergelijke beschuldigingen.

De wrakingsprocedure: sub-discussie en specifiek activiteitstype

Het recht van burgers om toegang te hebben tot een onafhankelijke rechter is vastgelegd in internationale, Europese en nationale wet- en regelgeving. Als rechters zelf voorzien dat er twijfels zouden kunnen rijzen omtrent hun onpartijdigheid, kunnen zij zich verschonen. Een reden voor een verschoningsverzoek kan liggen in een persoonlijke verhouding met (een van) de procespartijen of advocaten. Als de rechter zelf geen aanleiding ziet zich te verschonen, maar een van de procespartijen twijfelt aan de onpartijdigheid van de rechter aan wie de zaak is toegewezen dan kan die partij een wrakingsprocedure in gang zetten. De wrakingskamer die over wrakingsverzoeken beslist bestaat uit rechters bij de rechtbank waar het oorspronkelijke proces is gestart.

Om de argumentatie van procespartijen en van de wrakingskamer systematisch te identificeren en kritisch te beoordelen, maak ik onder andere gebruik van het pragma-dialectische onderzoek naar juridische argumentatie dat is gestart door, onder anderen, Feteris (1989), Kloosterhuis (2002) en Plug (2000). Een belangrijk deel van dit onderzoek is erop gericht na te gaan hoe de institutionele kenmerken van juridische activiteitstypen, zoals een juridische procedure, bepalend zijn voor de argumentatie die in juridische contexten naar voren wordt gebracht. Resultaten van dit onderzoek vormen het vertrekpunt voor de analyse en evaluatie van argumentatie die naar voren wordt gebracht in wrakingsprocedures.

De wrakingsprocedure kan worden gereconstrueerd als een sub-discussie die is gesitueerd in de openingsfase van de discussie over het initiële, juridische geschil: de hoofddiscussie. In de sub-discussie gaat het verschil

van mening over of de rechter al dan niet aan de discussie over het initiële geschil mag blijven deelnemen. Nadat het verschil van mening daarover in de sub-discussie is opgelost, kan de hoofddiscussie worden voortgezet.

Aangezien de wrakingsprocedure eigen institutionele randvoorwaarden kent, moet de wrakingsprocedure worden beschouwd als een subtype van het argumentatieve activiteitstype juridische procedures (Plug, 2016b, p. 208). In deze procedure heeft het geschil standaard betrekking op de propositie 'Het verzoek om rechter (X) te wraken moet worden toegewezen'. Ook de verdeling van de rollen is institutioneel vastgelegd. De procespartij die de onpartijdigheid van de rechter in twijfel trekt, neemt de positie in van protagonist. De rol van de rechter wiens onpartijdigheid wordt betwist kan worden gereconstrueerd als die van een antagonist of protagonist. De rol van de wrakingskamer is die van de (institutionele) antagonist.¹

Vervolgens bepalen de institutionele randvoorwaarden van het activiteitstype ook voor een deel de argumentatie die aan een wrakingsverzoek ten grondslag moet liggen. De algemene structuur bestaande uit expliciete en impliciete argumenten wordt bepaald door zowel relevante wetgeving als door jurisprudentie. De criteria die zijn bepaald in wetgeving zijn vastgelegd in art. 512 Sv, art. 36 Rv en art. 8:15 Awb. De belangrijkste criteria in de jurisprudentie zijn ontwikkeld in de zogenaamde Hauschildt-zaak door het Europese Hof voor de Rechten van de Mens (ECHR, 24 Mei 1989, *NJ* 1990, 627). In de uitspraak over die zaak wordt een onderscheid gemaakt tussen een subjectieve en een objectieve test. De objectieve test bepaalt zich tot de omstandigheden van de zaak die, nog afgezien van het persoonlijke gedrag van de rechter op de zitting, tot vrees voor partijdigheid aanleiding kunnen zijn. Dit is bijvoorbeeld het geval als een rechter een (persoonlijk) belang heeft in een zaak of eerder in een andere hoedanigheid betrokken was bij de zaak waarover hij moet vonnissen.

De subjectieve test heeft betrekking op de persoonlijke instelling en houding van de rechter in de voorliggende zaak of ten opzichte van de bij de zaak betrokken partijen. De houding van een rechter of zijn of haar bevooroordeelde instelling kunnen worden afgeleid van zijn of haar verbale gedrag gedurende rechtszittingen.

Argumenten in een sub-discussie over het verbale gedrag van de rechter

Kritiek op het verbale gedrag van de rechter tijdens de rechtszitting is een veel voorkomende reden om een wrakingsverzoek in te dienen (Knapen, 2012; Hammerstein, 2014). De kritiek kan betrekking hebben op opmerkingen die de indruk wekken dat de rechter al op voorhand overtuigd is van

de schuld van een procespartij of op een informele woordkeus die lijkt te duiden op sympathie voor het standpunt van een van de partijen.² Of het verbale gedrag van een rechter in een bepaalde zaak al dan niet voldoet aan de voorwaarde van onpartijdigheid maakt onderdeel uit van de subjectieve test. Bij de beoordeling daarvan moet worden uitgegaan van de zogenaamde persoonlijke onpartijdigheid van een rechter totdat het tegendeel wordt aangetoond. Op grond van argumenten moet dus worden aangetoond dat het vermoeden van vooringenomenheid van de rechter gerechtvaardigd is.

In het geval twijfel over de onpartijdigheid van de rechter is gebaseerd op zijn verbale gedrag, kan ervan worden uitgegaan dat de argumentatie bestaat uit, bijvoorbeeld, een citaat van de woordkeus van de rechter gedurende de rechtszitting of uit een verwijzing naar het gebruik door de rechter van retorische stijlmiddelen zoals vergelijkingen, hyperbolen of metaforen. Deze voorbeelden van verbaal gedrag kunnen worden beschouwd als specifieke eigenschappen die wijzen op een gebrek aan persoonlijke onpartijdigheid.

De argumentatiestructuur van de sub-discussie betreffende de wraking van een rechter op grond van zijn of haar verbale gedrag kan worden weergegeven als Figuur 1 op de volgende pagina.

In Figuur 1 is het wrakingsverzoek uitsluitend gebaseerd op klachten over het verbale gedrag van de rechter. In de juridische praktijk kan deze grond voor twijfel over juridische onpartijdigheid worden gecombineerd met andere klachten over het (verbale) gedrag van de rechter (Plug, in press). In deze gevallen bestaat de argumentatie die aan het wrakingsverzoek ten grondslag ligt uit complexe argumentatie.

Kritiek op het gebruik van retorische vragen

Een van de terugkerende klachten over het verbale optreden van de rechter op de zitting, heeft betrekking op het gebruik van retorische vragen door de rechter.³ In deze paragraaf ga ik na hoe deze klachten deel uit kunnen maken van de argumentatie die een procespartij aan een wrakingsverzoek ten grondslag legt. Vervolgens laat ik zien dat klachten over het gebruik van retorische vragen kunnen worden opgevat als beschuldigingen van ontsporingen van strategisch manoeuvreren. Ook laat ik zien dat deze klachten over het gebruik van retorische vragen door de rechter het gevaar lopen zelf eveneens beschouwd te worden als ontsporingen van strategisch manoeuvreren.

1.

Het verzoek om rechter (X) te wraken moet worden toegewezen

1.1

(1.1')

Rechter (X) is niet onpartijdig (Y) ----&----
in de onderhavige zaak

(Art.512 Sv/36 Rv/8:15 Awb)

1.1.1

(1.1.1')

(De vrees voor) subjectieve ----&----
partijdigheid van rechter (X) is
gerechtvaardigd (subjectieve test)

(ECHR, 24 Mei 1989, *NJ* 1990, 627)

1.1.1.1

Het verbale gedrag van rechter (X) tijdens
de behandeling van de zaak getuigt van een
gebrek aan (persoonlijke) onpartijdigheid (Y)

1.1.1.1.1

----&----

(1.1.1.1.1')

Rechter (X) heeft (Z)
naar voren gebracht

(Z) duidt op vooringenomenheid (Y)

Figuur 1 De argumentatiestructuur van een sub-discussie over onpartijdigheid gebaseerd op het verbale gedrag van de rechter

De beoordeling van kritiek op retorische vragen

De retorische vraag (Erotesis) wordt beschouwd als een vraag die geen antwoord behoeft, omdat het antwoord wordt geïmpliceerd of omdat het antwoord voor de hand liggend is in de gegeven context. De vraagvorm wordt niet gebruikt om een verzoek om informatie tot uitdrukking te brengen, maar om een bewering tot uitdrukking te brengen.

Studies van argumentatietheoretici naar de retorische vraag richten zich op de rol van de retorische vraag in argumentatieve contexten. In deze studies van onder andere Slot (1993), Ilie (1994, 1995), Van Eemeren, Houtlosser

en Snoeck Henkemans (2007) en Snoeck Henkemans (2006), worden drie argumentatieve hoofdfuncties van retorische vragen onderscheiden. De eerste functie is die van het naar voren brengen van een standpunt. De tweede functie is het gebruik van een retorische vraag om een argument naar voren te brengen. De derde functie is die van het gebruik van een retorische vraag om iets als een gemeenschappelijk uitgangspunt voor te stellen in de openingsfase van een discussie.

Uit literatuur waarin het gebruik van de retorische vraag in de rechtszaal wordt besproken, onder andere door Danet (1980), Pascual (2006) en Tiersma (2000), komt naar voren dat alle drie de argumentatieve functies in de praktijk van toepassing zijn, zij het dat ze niet als zodanig worden benoemd. Deze literatuur bespreekt echter niet hoe kritiek op de retorische vraag een rol kan spelen in een wrakingsprocedure. Vanuit een argumentatietheoretisch perspectief kan die kritiek worden opgevat als een beschuldiging van ontspoord strategisch manoeuvreren door de rechter. Aan de hand van de motivering van de uitspraak over het wrakingsverzoek kan worden gereconstrueerd of en zo ja hoe de wrakingskamer die beschuldiging beoordeelt.

De volgende wrakingszaak die werd behandeld door de wrakingskamer illustreert welke complicaties zich kunnen voordoen bij de beoordeling en de rechtvaardiging van klachten over het gebruik van retorische vragen door de rechter tijdens de zitting. Het verschil van mening in deze zaak betreft het recht van de verzoeker om de rechter te verzoeken zich strikt te houden aan de termijn van tien dagen uit artikel 8.58 van de Algemene wet bestuursrecht (Awb), hetgeen inhoudt dat documenten die te laat worden ingebracht niet aan het dossier worden toegevoegd. Het wrakingsverzoek heeft betrekking op een vraag van de rechter tijdens de rechtszitting over het verzoek om de termijn van tien dagen te handhaven. De klacht dat het hier een retorische vraag betreft wordt als volgt onderbouwd:

- (1) (...) Uit het proces-verbaal van de zitting volgt voorts dat de rechter naar aanleiding van het beroep van verzoeker op de tien-dagentermijn, zoals bedoeld in artikel 8:58 Algemene wet bestuursrecht (Awb), de vraag heeft gesteld “wat wilt u daarmee bereiken?”.

(Rechtbank Den Haag, 3 April 2015, ECLI:NL:RBDHA:2015:4564)

Uitgaand van de algemene argumentatiestructuur uit de tweede paragraaf kan de argumentatie waarmee de verzoeker het gebruik van de retorische vraag door de rechter bekritiseert, als volgt worden gereconstrueerd.

1.

Het verzoek om rechter (X) te wraken moet worden toegewezen

|
(1.1-1.1.1)

|
1.1.1.1

Het verbale gedrag van rechter (X) tijdens de zitting getuigt van een gebrek aan (persoonlijke) onpartijdigheid (Y)

|
1.1.1.1.1 ---&--- (1.1.1.1.1')

Rechter (X) heeft een retorische vraag naar voren gebracht (Z)

een retorische vraag (Z)
duidt op vooringenomenheid (Y)

|
1.1.1.1.1.1

“Wat wilt u daarmee bereiken?”

Figuur 2 De argumentatiestructuur van een sub-discussie over onpartijdigheid van de rechter, op grond van het gebruik van een retorische vraag

De wrakingskamer die het wrakingsverzoek moest beoordelen, honoreerde de klacht over het gebruik van de retorische vraag door rechter. De wrakingskamer argumenteerde langs ongeveer dezelfde lijnen als de partij die het wrakingsverzoek indiende.

(2) (...) Het is vervolgens aan de rechter om aan de hand van het criterium van de goede procesorde te beoordelen of de stukken daadwerkelijk worden toegelaten, waarbij de rechter kan vragen om het verzoek te motiveren. De wrakingskamer is van oordeel dat genoegzaam vaststaat dat de rechter dit heeft gedaan in de vorm van een retorische vraag, nu dit door de rechter niet is weersproken. Door deze vorm te bezigen kan ook op dit punt de indruk zijn ontstaan dat de rechter niet geheel onbevangen was ten aanzien van de zaak.

(Rechtbank Den Haag, 3 April 2015, ECLI:NL:RBDHA:2015:4564)

De rechtvaardiging van kritiek op het gebruik van retorische vragen

In de voorbeelden (1) en (2) is de relatie tussen het argument en het (sub) standpunt gebaseerd op symptomatische argumentatie. In dit type argumentatie wordt hetgeen in het argument wordt gezegd gezien als een symptoom, een indicatie of een kenmerk van wat wordt beweerd in het (sub)standpunt. Met andere woorden, het naar voren brengen door de rechter van een retorische vraag wordt gezien als een indicatie dat hij niet onpartijdig is. Het argumentatieschema voor symptomatische argumentatie (Van Eemeren et al. 2007, p. 154) kan als volgt worden toegepast op de argumentatie in het voorbeeld.

Standpunt:	Voor X geldt Y Voor rechter (X) geldt een gebrek aan (persoonlijke) onpartijdigheid (Y)
Want:	Voor X geldt Z Voor rechter (X) geldt dat hij retorische vraag naar voren heeft gebracht (Z)
En:	Z is kenmerkend voor Y Het naar voren brengen van een retorische vraag (Z) duidt op vooringenomenheid (Y)

Dit schema kan worden gebruikt als uitgangspunt voor de beoordeling van de argumentatie: het laat zien op welke twee premissen de beoordeling zich moet concentreren. Voor de beoordeling van de eerste premisse moet nagegaan worden of een bepaalde uitspraak van de rechter inderdaad kan worden opgevat als retorische vraag. De rechtvaardiging die wordt gegeven in voorbeeld (1) beperkt zich tot een citaat van de vraag die tijdens de rechtszitting door de rechter werd gesteld. Deze verwijzing naar de precieze formulering van de vraag is niet zonder belang omdat, zoals hier oorspronkelijk ook het geval was, de rechter onjuist zou kunnen zijn geciteerd. Het citaat, dat op zichzelf eveneens kan worden beschouwd als symptomatische argumentatie, is echter geen voldoende rechtvaardiging voor de conclusie dat de vraag van de rechter als een retorische vraag moet worden opgevat.

Ten eerste impliceert het argument ten onrechte dat een vraag als zodanig ondubbelzinnig als retorisch kan worden gekwalificeerd. Uit onderzoek van bijvoorbeeld Frank (1990), Ilie (1994) en Slot (1993) naar de identificatie van retorische vragen blijkt echter dat zo'n eenduidige interpretatie niet mogelijk is. Een argument dat uitsluitend bestaat uit een citaat van de vraag in kwestie, zonder dat daar een aanvullende rechtvaardiging aan is toegevoegd, kan daarom worden opgevat als een drogredelijke, ontspoorde

discussiezet. Door de indruk te wekken de vraag van de rechter opzettelijk onjuist te interpreteren, ontspoord de strategische manoeuvre als gevolg van een overtreding van de Taalgebruiksregel die onder andere bepaalt dat ‘discussianten de formuleringen van de andere partij niet opzettelijk onjuist mogen interpreteren.’

Ten tweede laat het argument, zoals geformuleerd in de eerste premisse, impliciet welke indirecte taalhandeling met de retorische vraag tot uitdrukking wordt gebracht. Gelet op de institutionele karakteristieken van de wrakingsprocedure, is het van belang vast te stellen of de retorische vraag in kwestie kan worden opgevat als een beweeder waarmee in principe een standpunt tot uitdrukking zou kunnen zijn gebracht. Gegeven dat de rol van de rechter in de initiële discussie die van institutionele antagonist is, zou in de wrakingsprocedure moeten worden gerechtvaardigd dat de rechter desondanks gehouden kan worden aan het naar voren brengen van een beweeder waarmee hij een standpunt tot uitdrukking brengt.

De analyse van de complicaties betreffende de eerste premisse in het argumentatieschema laat zien dat de argumentatie die naar voren is gebracht ter rechtvaardiging van de stelling dat de uitspraak van de rechter als een retorische vraag zou kunnen worden beschouwd, moet worden uitgebreid. Als vertrekpunt voor deze uitbreiding kan gebruik worden gemaakt van de benadering van Snoeck Henkemans (2006). Deze benadering voor de analyse van retorische vragen is gebaseerd op de taalhandelingstheorie. In deze benadering wordt een retorische vraag beschouwd als een indirecte taalhandeling die een beweeder tot uitdrukking brengt omdat met deze taalhandeling, als we haar letterlijk nemen, twee regels voor communicatie worden overtreden. In de eerste plaats kent degene die de vraag stelt het antwoord erop al, hetgeen de vraag overbodig maakt. In de tweede plaats is de vraag onoprecht aangezien degene die de vraag stelt geen antwoord van de aangesproken gesprekspartner verwacht. Gebaseerd op deze criteria kan een voorstel voor uitbreiding (cursief) van de argumentatie aangaande de eerste propositie in het schema er als volgt uitzien:

- 1.1.1.1.1 Rechter (X) heeft een retorische vraag (Z) naar voren gebracht
- 1.1.1.1.1a Rechter (X) heeft vraag (V) gesteld
- 1.1.1.1.1b *Vraag (V) moet geïnterpreteerd worden als een beweeder*
 - 1.1.1.1.1b.1a *De letterlijke vraag (V) is overbodig*
 - 1.1.1.1.1b.1b *De letterlijke vraag (V) is onoprecht*
 - 1.1.1.1.1b.1c *De interpretatie van (V) als een beweeder heft de overbodigheid en de onoprechtheid van (Z) op*
- (1.1.1.1.1') RV (Z) duidt op vooringenomenheid (Y)

De beoordeling van de tweede premisse is gericht op de rechtvaardigende kracht van het argument dat wordt uitgedrukt in (1.1.1.1.1'): De retorische vraag van de rechter duidt op vooringenomenheid. In dit geval concentreren de kritische vragen die betrekking hebben op het argumentatieschema zich op de symptomatische relatie waarop de argumentatie is gebaseerd. De relevante vraag is of het gebruik van een retorische vraag inderdaad op vooringenomenheid duidt. Als we de institutionele kenmerken van een juridische procedure in aanmerking nemen, hoeft het gebruik van een retorische vraag niet per se te duiden op vooringenomenheid. Dat is wel het geval als de retorische vraag kan worden opgevat als een negatief standpunt ten opzichte van (de positie van een van) de procespartijen in de onderhavige zaak. Om die reden zou de argumentatie, anticiperend op de kritische vraag met betrekking tot de symptomatische relatie tussen het argument en het (sub-)standpunt, als volgt kunnen worden uitgebreid:

- (1.1.1.1.1') RV (Z) duidt op vooringenomenheid (Y)
- (1.1.1.1.1':1a) *RV (Z) functioneert als een beweerder waarmee een (negatief) standpunt van rechter (X) tot uitdrukking wordt gebracht*
- (1.1.1.1.1':1b) *het (negatieve) standpunt van rechter (X) heeft betrekking op de zaak waar hij over moet oordelen*
- (1.1.1.1.1':1a/b') *een beweerder waarmee een negatief standpunt van rechter (X) tot uitdrukking wordt gebracht dat betrekking heeft op de zaak waar hij over moet oordelen, duidt op partijdigheid (Y)*

Dit argumentatieve overzicht verschaft een model dat als vertrekpunt kan dienen voor de beoordeling van de argumentatie in wrakingsverzoeken die betrekking hebben op het gebruik van retorische vragen door de rechter. Het argumentatieve overzicht is gebaseerd op voorwaarden die zijn ontleend aan (argumentatietheoretisch) onderzoek naar retorische vragen en vormt een uitbreiding van het argumentatieve overzicht dat is gebaseerd op de juridische voorwaarden voor het toekennen van een wrakingsverzoek (Figuur 1 en Figuur 2). Het model heeft als doel meer inzicht te geven in de reikwijdte van de argumentatie die noodzakelijk is voor de onderbouwing van interpretatieve standpunten ten aanzien van (vermeende) retorische vragen. Deze inzichten kunnen behulpzaam zijn bij zowel de presentatie als bij de beoordeling van de argumentatie die ten grondslag ligt aan het soort klachten dat betrekking heeft op het verbale gedrag van de rechter.

Aan de hand van het model kan ten aanzien van de voorbeelden (1) en (2) worden vastgesteld dat de argumentatie van zowel de verzoeker tot wraking als die van de wrakingskamer tekort schiet. Dat betekent in

pragma-dialectische termen, dat het strategisch manoeuvreren van de verzoeker en de wrakingskamer in deze voorbeelden ontspoord doordat de argumentatie zoals gepresenteerd in het model, ontbreekt. Allereerst ontspoord het strategisch manoeuvreren doordat zonder uitbreiding van de argumentatie, de bij het kentekenschema behorende kritische vraag ‘of de geciteerde vraag niet ook als een niet-retorische vraag kan worden opgevat,’ niet bevredigend beantwoord kan worden. Omdat daarmee de indruk gewekt kan worden dat de vraag van de rechter opzettelijk onjuist is geïnterpreteerd, kan de ontsporing ook het gevolg zijn van een overtreding van de Taalgebruiksregel. Tot slot ontspoord het strategisch manoeuvreren doordat het naar voren brengen van een retorische vraag ten onrechte als een voldoende voorwaarde wordt gepresenteerd om te concluderen dat die vraag op vooringenomenheid duidt.

Deze ontsporingen hebben niet alleen tot gevolg dat het onderhavige geschil niet op een aanvaardbare manier is opgelost. De ontsporingen in met name het strategisch manoeuvreren van de wrakingskamer brengen ook het risico met zich mee dat vragen die de rechter tijdens de behandeling van een zaak stelt, te gemakkelijk door partijen worden aangegrepen als grond voor een wrakingsverzoek.⁴

Conclusie

In deze bijdrage heb ik verslag gedaan van onderzoek naar klachten die naar voren worden gebracht in wrakingsprocedures over het verbale optreden van de rechter tijdens de zitting. Ik heb laten zien hoe op basis van institutionele randvoorwaarden voor deze procedures zoals die zijn geformuleerd in wetgeving en jurisprudentie, een algemene argumentatiestructuur opgesteld kan worden voor de argumentatie die naar voren moet worden gebracht om een wrakingsverzoek dat betrekking heeft op het verbale gedrag van de rechter te rechtvaardigen. Deze structuur vormt het uitgangspunt voor de analyse en beoordeling van kritiek die specifiek gericht is op het gebruik van retorische vragen door de rechter. Vanuit een pragma-dialectische perspectief komt deze kritiek erop neer dat de rechter van ontspoord strategisch manoeuvreren wordt beschuldigd. Aan de hand van een voorbeeld heb ik laten zien dat de kritiek van de procespartij of de wrakingskamer op de rechter zelf ook weer tot een ontsporing van strategisch manoeuvreren kan leiden als de kritiek onvoldoende is onderbouwd. Om die ontsporingen te helpen voorkomen heb ik een uitbreiding voorgesteld van de algemene argumentatiestructuur voor wrakingsverzoeken. Deze

uitbreiding beoogt aanknopingspunten te bieden voor zowel de rechtvaardiging van de klachten van procespartijen over het gebruik van retorische vragen door de rechter als voor de rechtvaardiging van de beoordeling van die klachten door de wrakingskamer.

Noten

1. In haar door het recht bepaalde discussierol, trekt de wrakingskamer (impliciet) het standpunt van de verzoeker tot wraking in twijfel alvorens na te gaan of de verzoeker erin slaagt die twijfel weg te nemen (zie ook Feteris 1989, p. 119).
2. In een toelichting op het wrakingsverzoek dat Wilders indiende op 3 november 2016 wijst zijn advocaat, Knoops, bijvoorbeeld op het gebruik door de rechter van het adjectief 'heel'. In het getuigenverhoor noemde de rechter iets niet 'strikt' maar 'heel strikt'. Aangezien de rechter daarnaast de getuige ook 'bot' onderbrak, concludeert Knoops: 'Dit is niet de schijn van partijdigheid, dit is gewoon partijdigheid' (*Volkskrant*, 4 november 2016).
3. Zie bijvoorbeeld: Rechtbank Gelderland, 8 Februari 2018, ECLI:NL: 2018:630, Rechtbank Zeeland-West-Brabant, 31 Mei 2016, ECLI:NL:RBZWB:2016:3618, Rechtbank Amsterdam, 1 December 2016, ECLI: RBAMS:2016:8963, Rechtbank Overijssel, 12 Juni 2014, ECLI:NL:RBOVE:2014:3237, Rechtbank Gelderland, 15 Augustus 2013, ECLI:NL:RBGEL:2013:3766, Rechtbank Gelderland, 30 November 2010, ECLI:NL: 2010:BO6401.
4. Zie Giesen, Kristen, Enneking, & Van Lent (2013) voor onderzoek naar het oneigenlijke gebruik van de wrakingsprocedure.

Bibliografie

- Bauw, E. (2011). Wat te denken van wraking? *Ars Aequi*, 60(3), 202-206.
- Danet, B. (1980). Language in the legal process. *Law & Society Review*, 14(3), 445-564.
- Eemeren, F.H. van (2010). *Strategic Maneuvering in Argumentative Discourse. Extending the Pragma-dialectical Theory of Argumentation*. Amsterdam/Philadelphia, PA: John Benjamins.
- Eemeren, F.H. van, Houtlosser, P., & Snoeck Henkemans, A.F. (2007). *Argumentative indicators in discourse. A pragma-dialectical study*. Dordrecht: Springer.
- Feteris, E.T. (1989). *Discussieregels in het recht. Een pragma-dialectische analyse van het burgerlijk proces en het strafproces*. Dordrecht: Foris.
- Frank, J. (1990). You call that a rhetorical question? Forms and functions of rhetorical questions in conversations. *Journal of Pragmatics*, 14(5), 723-738.
- Giesen, I, Kristen, F., Enneking, L., & Lent, L. van. (2013). Op weg naar een nieuwe wrakingsprocedure. Meer legitimiteit en minder oneigenlijk gebruik. (Challenging Judges: A New Procedure for Disqualification of Judges in the Netherlands?). *Nederlands Juristenblad*, 8, 466- 477.
- Hammerstein, A. (2014). Onpartijdigheid in het geding. *Trema*, 26(5), 148-154.
- Ilie, C. (1994). *What else can I tell you? A pragmatic study of English rhetorical questions as discursive and argumentative acts*. PhD dissertation, University of Stockholm.

- Ilie, C. (1995). The validity of rhetorical questions as arguments in the courtroom. In F.H. van Eemeren, R. Grootendorst, J.A. Blair, & C.A. Willard (Eds.) *Proceedings of the Third ISSA Conference on Argumentation. Special Fields and Cases. Volume 5*, (pp.73-88). Amsterdam: Sic Sat.
- Kloosterhuis, H.T.M. (2002). *Van overeenkomstige toepassing: de pragma-dialectische reconstructie van analogie-argumentatie in rechterlijke uitspraken*. Amsterdam: Thela Thesis.
- Knapen, M. (2012). Advocaten ontdekken wraking. *Advocatenblad*, nr. 1, 18-23.
- Pascual, E. (2006). Questions in legal monologues: Fictive interaction as argumentative strategy in a murder trial. *Text & Talk*, 26(3), 10-21.
- Plug, H.J. (2000). *In onderlinge samenhang bezien. De pragma- dialectische reconstructie van complexe argumentatie in rechterlijke uitspraken*. Amsterdam: Thela Thesis.
- Plug, H.J. (2016a). Administrative judicial decisions as a hybrid argumentative activity type. *Informal Logic*, 36 (3), 333-348.
- Plug, H.J. (2016b). Strategisch manoeuvreren met 'role shifts' in bestuursrechtelijke uitspraken: de complexe rol van de rechter. In D. van de Mierop, L. Buysse, R. Coesemans, & P. Gillaerts (Red.), *De macht van de taal: Taalbeheersingsonderzoek in Nederland en Vlaanderen* (pp. 205-218). Leuven/Den Haag: Acco.
- Plug, H.J. (in press). The analysis of argumentation underlying complaints about a lack of judicial impartiality. In M. Manzin, F. Puppò & S. Tomasi (Eds.). *Studies on Argumentation & Legal Philosophy. Multimodal Argumentation, Pluralism and Images in Law*. Trento: Quaderni della Facoltà, Università di Trento.
- Rossum, W. van, Tigchelaar, J., & Ippel, P. (2012). *Wraking bottom-up. Een empirisch onderzoek*. Den Haag: Raad voor de Rechtspraak.
- Slot, P. (1993). *How can you say that? Rhetorical questions in argumentative texts*. Amsterdam: IFOTT.
- Snoeck Henkemans, A.F. (2006). Manoeuvring strategically with rhetorical questions. In F. H. van Eemeren, J.A. Blair, C.A. Willard & B. Garssen (Eds.) *Proceedings of the Sixth Conference of the International Society for the Study of Argumentation (ISSA)* (pp. 1-11). Amsterdam: Rozenberg.
- Tiersma, P.M. (2000). *Legal Language*. Chicago, IL: University of Chicago Press.

De redacteur als tekstbeoordelaar

Acquisitie en redactie bij Nederlandse algemene uitgeverijen

Everdien Rietstap
Universiteit van Amsterdam

TVT 41 (1): 265–280

DOI: 10.5117/TVT2019.1.018.RIET

Abstract

The Editor as Text Evaluator: Acquisition and Editing at Dutch General Publishing Companies

In this article the contribution of the editor in the book production process of Dutch general publishing companies is explored. Editors are involved in and responsible for text evaluation in two stages of the book production process. Results from book publishing studies show the underlying criteria for the acquisition of new books. The editor evaluates whether a new work is saleable, important for the identity of the publishing company, based on agreements with the author, and editable. When the book proposal or manuscript has been accepted by the publisher the editing process starts. In this institutional context a good relation between editor and author is important. The editor is of service to the author and the book. By his knowledge of texts he should motivate and convince the author to rewrite his text, and in most cases has to accept it when the author rejects the proposed changes. By using the CCC-model Renkema, (2002; 2012) as an analytical instrument for text genetics, we see that editors propose changes when the aim of the text does not correspond with the needs of the intended reader, if the text is not consistent, or if the text is not correct.

Keywords: publishing, book production, editing, text evaluation, tasks editor

Inleiding

Bij uitgeverijen speelt tekstbeoordeling een belangrijke rol, zowel bij de acquisitie van nieuw werk als tijdens het redactieproces. Tijdens de transformatie van manuscript naar boek werken redacteur en auteur nauw met elkaar samen: de redacteur evalueert de tekst op verschillende momenten en verstrekt feedback, en de auteur herschrijft de tekst, en accepteert al dan niet de voorgestelde wijzigingen. Cultuursocioloog Howard Becker (2008) beschouwt een boek daarom als product van een gezamenlijke inspanning, waarbij naast de schrijver zelf ook anderen binnen (en buiten) de uitgeverij betrokken zijn.

Hoewel over acquisitie en redactie bij Nederlandse algemene uitgeverijen nog nauwelijks is gepubliceerd (De Glas, 2006, p. 8), vormt het redactieproces het hart van de uitgeverij. Door de onderliggende criteria bloot te leggen, ontstaat een beeld van de inhoud van het redacteurschap bij uitgeverijen die fictie, non-fictie en/of kinder- en jeugdboeken uitgeven. Volgens Coser, Kadushin en Powell (1985, p. 124) zijn redacteurs niet in staat om systematisch te beschrijven hoe ze manuscripten verkrijgen en rekenschap te geven van de factoren die van invloed zijn op het besluitvormingsproces. Coser et al. tekenen hierbij aan dat ervaren redacteurs minder geneigd zijn om op hun intuïtie te varen dan minder ervaren redacteurs.

Dit exploratieve onderzoek heeft als doel om de criteria bijeen te brengen die redacteurs hanteren bij de acquisitie en redactie van oorspronkelijk Nederlandstalige boeken bij Nederlandse algemene uitgeverijen. Het richt zich enerzijds op de criteria die een rol spelen bij het besluitvormingsproces rondom het verwerven en beoordelen van nieuw werk en anderzijds op de criteria die een rol spelen bij het geven van feedback, het redigeren en corrigeren.

Het accent van het onderzoek ligt op de beoordeling van teksten; het onderzoek illustreert op welke wijze de redacteur bijdraagt aan de productie van boeken. Met behulp van wetenschappelijke literatuur op het snijvlak van boekwetenschap en taalbeheersing worden de criteria besproken die ten grondslag liggen aan acquisitie en redactie. In de praktijk ontbreken echter harde gegevens. Daarom wordt aan de hand van tekstgenetisch onderzoek bekeken of de hierboven aangeduide criteria terug te zien zijn in de praktijk van de Nederlandse uitgeverij. Voor de analyse van de redactiecriteriën wordt gebruikgemaakt van het CCC-model van Renkema (2002, p. 181; 2012, p. 30). Hoewel dit model oorspronkelijk is ontworpen voor de analyse en evaluatie van zakelijke teksten, blijkt het ook bruikbaar om op

systematische wijze redactievoorstellen te verklaren tegen het licht van genre en doelgroep.

Aan de hand van de drie C's – correspondentie, consistentie en correctheid – kan worden geanalyseerd hoe de verstrekte feedback en wijzigingsvoorstellen te relateren zijn aan de opvatting over tekstkwaliteit (Van de Poel, Carstens, & Linnegar, 2012). Het is complex om casestudies te vinden. Zowel redacteuren als schrijvers gooien de verschillende versies van manuscripten en de verstrekte feedback vaak weg nadat het boek is verschenen (Buschenhenke, z.d.). Voor dit onderzoek is daarom gebruikgemaakt van vijftig beschikbare productiedossiers, becommentarieerde manuscripten en correspondentie uit het archief van het Literatuurmuseum, en uitgevers- en schrijversarchieven.¹

Het onderzoek sluit aan bij de door Janssen (2000, p. 66) gesignaleerde wetenschappelijke belangstelling naar de omstandigheden en voorwaarden waaronder literatuur binnen uitgeverijen wordt gecreëerd. In haar bijdrage 'Onderzoek naar twintigste-eeuwse uitgeverijen. Een stand van zaken' bepleit ze nader onderzoek naar het functioneren van literaire uitgeverijen die de theorievorming ten goede komt. Janssen benadrukt dat in dit onderzoek bij voorkeur 'een bepaald aspect van het functioneren van literaire uitgeverijen op systematische wijze onder de loep wordt genomen' (Janssen, 2000, p. 78), liefst in vergelijking met andere algemene uitgeverijen en in historisch perspectief. Hoewel er monografieën (onder meer *Text editing* van Van de Poel, Carstens, & Linnegar, 2012 en *Editing Fact and Fiction. A Concise Guide to Book Editing* van Sharpe & Gunter, 1994) en praktische handboeken (onder meer *Handboek voor de redacteur* van Aalbrecht & Wagenaar, 2015) zijn geschreven over redactie, ontbreekt systematisch onderzoek.

Het productieproces van boeken

Voordat we gaan inzoomen op tekstevaluatie en -redactie bij de algemene uitgeverij is het goed een beknopt beeld te geven van de institutionele context waarbinnen de tekstbeoordeling plaatsvindt. Nederlandse uitgevers van fictie, non-fictie en kinder- en jeugdboeken hebben in 2016 132.000 verschillende Nederlandstalige boektitels verkocht. 95% van deze titels is uitgegeven door 157 beeldbepalende uitgeverijen (KVB Boekwerk, z.d.). Binnen deze uitgeverijen vervullen de uitgever, de redacteur en de auteur verschillende rollen; de benamingen zijn echter niet bij alle uitgeverijen hetzelfde. In deze bijdrage staat de uitgever voor degenen die bepaalt welke

boeken op de markt worden gebracht en de redacteur voor degene die acquireert en verantwoordelijk is voor het redactieproces.

De uitgever is als cultureel ondernemer zakelijk en artistiek leider ineen: hij is verantwoordelijk voor een goede bedrijfsvoering, maar levert via zijn uitgeefbeleid ook een bijdrage aan de Nederlandse boekencultuur. Hij bepaalt welke boeken gepubliceerd worden en welke niet, en is verantwoordelijk voor de fondsvorming die de uitgeverij een gezicht geeft en onderscheidt van de concurrentie (Kuitert, 2008a). Socioloog Lewis Coser kent uitgeverijen daarom de rol toe van poortwachters van ideeën: ‘gatekeepers of ideas in as much as they are empowered to make decisions as to what is let “in” and what is “kept out”’ (Coser, 1975, p. 15). De uitgever bepaalt immers voor welke titels en auteurs de deur opengaat en voor welke niet.

De uitgever werkt nauw samen met de redacteur, ook wel acquirerend redacteur of fondsredacteur genoemd. De redacteur onderhoudt contact met de schrijvers uit zijn fonds, beoordeelt binnengekomen boekvoorstellen en manuscripten, schrijft leesrapporten en gaat actief op zoek naar nieuwe titels en auteurs: de redacteur is verantwoordelijk voor de acquisitie van nieuw werk dat aansluit bij het uitgeefbeleid (Lewin, 1951, pp. 186-187). Wanneer een redacteur een manuscript de moeite waard vindt of een boekvoorstel inbrengt, moet de uitgever nog worden overtuigd om het boek uit te geven. Als de uitgever instemt met uitgave en de auteur een contract in het voortuitzicht stelt, valt de redactie van het manuscript onder de verantwoordelijkheid van de redacteur.

De schrijver wordt als ‘(literair) auteur’ beschouwd, zodra de uitgeverij zijn werk heeft geaccepteerd. Het feit dat de uitgeverij zijn werk heeft geaccepteerd, fungeert als kwaliteitskeurmerk en geldt bijvoorbeeld als een voorwaarde om subsidies te verkrijgen. Om de kwaliteit van het werk en het schrijfproces te bewaken, ontvangt de auteur redactionele begeleiding tijdens de verschillende fasen van het boekproductieproces. Artikel 3, lid 3 van het Modelcontract voor de uitgave van oorspronkelijk Nederlandstalig literair werk vermeldt hierover het volgende:

Indien de uitgever de kopij niet accepteert omdat de kwaliteit van de inhoud niet aan redelijke verwachtingen voldoet, kan de uitgever besluiten om niet tot uitgave over te gaan. Indien de uitgever de kopij accepteert, behoort het tot de taak van de uitgever om zo nodig, op kosten van de uitgever, de kopij in overleg met de auteur te redigeren. [...] (http://www.vvl.nu/vvl/media/original/59/modelcontract_oorspr_ned_lit_werk.pdf)

De inhoud van dit artikel legt een spanningsveld bloot. Hoewel de uitgever de kopij mag weigeren als die niet aan de verwachtingen voldoet en de redacteur de kopij in overleg met de auteur redigeert, neemt de auteur de voorstellen van de redacteur niet klakkeloos over. Wanneer er geen zwaarwegende redenen zijn om werk te weigeren, heeft de auteur in de regel het laatste woord. Zo volgde Rudy Kousbroek in zijn essaybundel *Medereizigers: over de liefde tussen mensen en dieren* (2009) de spellingregels van 1954 en weigerde hij zich te conformeren aan de spellingherziening uit 1995. Alle betrokken redacteurs van Uitgeverij Augustus wisten dit en hielden hier rekening mee (Master Redacteur/editor, 2017, p. 54).

Criteria acquisitie

Zoals we hebben gezien, zijn acquisitie en redactie twee verschillende activiteiten op de uitgeverij. Bij beide taken worden eigen criteria gehanteerd. Redacteurs nemen bij de acquisitie van nieuw werk het uitgeefbeleid van de uitgeverij als uitgangspunt. Er zijn verschillende studies verschenen die inzicht geven in de fondsvorming en het uitgeefbeleid van Nederlandse literaire uitgeverijen, zoals biografieën over uitgeverijen en hun uitgevers (onder meer de dissertatie *Geert van Oorschot, uitgever* van Fortuin, 2015), correspondentie tussen uitgevers en schrijvers, (auto)biografieën van schrijvers (onder meer *Dagelijks werk. Een schrijversleven* van Dorrestein, 2018), en studies naar de uitgeefbranche (onder meer *How Books Travel. Translation Flows and Practices of Dutch Acquiring Editors and New York Literary Scouts, 1980-2009* van Franssen, 2015).

Het beeld dat uit deze studies oprijst is dat uitgeefbeleid geen 'beleid' is, maar veelal van ad-hocbeslissingen aan elkaar hangt. Redacteurs lijken niet volgens vaste richtlijnen te werken, maar gebruiken ook iets ongrijpbaars als *fingerspitzengefühl*. Renkema bevestigt tijdens een interview dat hier twee andere, minder makkelijk te duiden C's betrokken worden in de tekstbeoordeling: de kwaliteit van het Concept en de Creativiteit (Master Redacteur/editor, 2018, p. 19).

Bij acquisitie speelt het aantal uit te geven boeken een rol. Binnen het uitgeefbeleid zijn twee uitersten te signaleren bij het nemen van uitgeefbeslissingen: de hagelschotmethode en *slow publishing* (Coser, 1975, p. 21). Bij de hagelschotmethode worden zo veel mogelijk boeken uitgegeven die passen binnen het fonds. Redacteurs streven ernaar om minimaal de gemaakte kosten terug te verdienen, zodat de continuïteit van de uitgeverij kan

worden gegarandeerd, en gaan ervan uit dat een aantal boeken resulteert in een verkoopsucces. In Nederland wordt deze methode door Uitgeverij Prometheus gebezigd: de uitgeverij heeft in 2018 180 titels op de markt gebracht in de fondsen kwalitatieve literaire fictie en diepgaande non-fictie (www.uitgeverijprometheus.nl). Bij de *slow-publishing* methode worden bewust minder boeken uitgegeven, een methode waarmee uitgeverij Das Mag experimenteert: de uitgeverij heeft sinds de oprichting in 2015 gemiddeld 10 titels per jaar uitgegeven. In de praktijk schommelen de meeste uitgeverijen tussen deze uitersten: ze doen aan risicospreiding én schenken extra tijd, geld en aandacht aan A-titels met veel potentie, zodat hun culturele en financiële continuïteit wordt gewaarborgd (Franssen, 2015).

Om hun acquisitietaak goed uit te kunnen voeren moeten redacteurs op de hoogte zijn van wat er speelt in het culturele veld en aftasten in welke genres, onderwerpen en auteurs de lezers geïnteresseerd zijn. Redacteurs putten in hun queeste naar nieuwe boeken uit hun netwerk en gaan actief op zoek naar schrijftalent via literaire en culturele festivals en evenementen, kranten, tijdschriften, (video)blogs, en schrijfpoleidingen en -wedstrijden. Daarnaast krijgen redacteurs de laatste twintig jaar via literair agenten en scouts steeds vaker manuscripten aangedragen en bekijken ze nog steeds de stapel ongevraagd toegestuurde manuscripten. Hoewel volgens Sander Blom (uitgever fictie bij Atlas Contact) minder dan één procent van de ingezonden manuscripten wordt uitgegeven, levert deze ene procent soms wel een bestseller op. Zo is werk van bekende, prijswinnende auteurs Dimitri Verhulst en Gustaaf Peek door respectievelijk door Sander Blom en Josje Kraamer (acquirerend redacteur bij Querido) uit de *slush pile* gehaald.

Uit de bestudeerde literatuur komen vier criteria naar voren die redacteurs hanteren bij het beoordelen van boekvoorstellen en manuscripten en bij het initiëren van nieuwe boeken. Ze stellen allereerst vast of het uit te geven boek verkoopbaar is. De originaliteit van het boek – de *unique selling points* –, de omvang van de doelgroep, maar ook het cv van de auteur spelen hierbij een belangrijke rol. Zo verkocht Meulenhoff ruim 20.000 exemplaren van *Gedichten van de broer van Roos* van de populaire BNN-presentator, blogger, columnist en dichter Tim Hofman, waar toonaangevende dichters veelal genoeg moeten nemen met enkele honderden verkochte exemplaren. Dit financiële criterium is over het algemeen van doorslaggevend belang, hoewel er binnen de uitgeverij een mate van persoonlijke stijl en beoordeling mag zijn die kenmerkend is voor culturele productie (Coser et al., 1985, pp. 144-146). Het is immers in de praktijk vaak lastig te voorspellen welk boek gaat verkopen en welk niet (Coser, 1975, p. 16).

Daarnaast beoordelen redacteuren of het boek past binnen de contouren van het fonds. Het uit te geven werk moet passen bij de identiteit van de uitgeverij en de uitgeeftraditie, en inhoudelijk meerwaarde hebben voor het fonds. Een debuut wordt vaak in een kleinere oplage uitgebracht, maar wanneer de auteur doorbreekt is deze investering goed voor de uitstraling en de financiële positie van de uitgeverij. Daarnaast stellen redacteuren soms boeken voor vanwege de status van het genre; dit geldt bijvoorbeeld voor poëziebundels die in de regel een lage oplage kennen. Binnen de uitgeverij worden deze investeringen gefinancierd met de uitgave van populaire genres zoals thrillers, *life style*- en kookboeken.

Het derde criterium betreft de gemaakte afspraken met de auteur. Op het moment dat de auteur zich aan de uitgeverij verbindt, is dit vaak niet eenmalig, al geldt voor veel debutanten dat zij geen groen licht krijgen voor volgende titels (Janssen, 2000, p. 68). Ook gevorderde schrijvers krijgen soms 'nee' te horen en stappen dan vaak over naar een andere uitgeverij.

Het vierde en laatste criterium betreft de investering in de redactie van het manuscript en de bereidheid van de auteur om zijn tekst te laten redigeren. Wat zijn de redactionele problemen, is de auteur nog bereid aan de tekst te werken én klikt het persoonlijk tussen redacteur en auteur? Uit interviews met redacteurs blijkt dat deze overwegingen worden meegenomen in hun besluitvormingsproces (Coser et al., 1985, p. 144). Dat de redacteur bereid is om te investeren in redactie als een werk potentie heeft, laat Kluun zien in 'The Directors Cut' (Kluun, z.d.). Op voorspraak van zijn uitgever Joost Nijsen (Podium) en redacteur sneuvelde bijna de helft van zijn ruim 150.000 woorden. 'Ja, die stukken zijn langdradig. Ja, ze zijn slecht geschreven. Vol spelfouten. Grammaticaal om te huilen. Compositorisch lachwekkend. Verhaaltechnisch volstrekt overbodig. Stijlistisch (sic) dramatisch: lees een willekeurig stuk en constateer dat de redacteur en meelezers gelijk hebben.' (Kluun, z.d.) Op basis van de eerste, ongeredigeerde versie van zijn manuscript wist hij echter wel een boekcontract voor zijn debuutroman *Komt een vrouw bij de dokter* binnen te slepen.

Redacteuren bouwen op basis van kennis en ervaring een referentiekader op dat het uitgangspunt vormt bij tekstbeoordeling. Ook reserveert de redacteur idealiter ruimte voor reflectie, zeker als een manuscript of boekvoorstel dat hij heeft afgewezen een succes wordt bij een concurrerende uitgeverij. Op het moment dat de redacteur groen licht krijgt van de uitgever gaat het redactieproces van start.

Criteria redactie

De essentie van het werk van een redacteur is dat deze zich dienstbaar opstelt bij de beoordeling van de tekst. Dienstbaar aan het boek, de lezer, de uitgeverij, en dienstbaar aan de auteur. In zoverre zijn de werkzaamheden van een redacteur te scharen onder de materiële productie. Kuitert (2008a) zet uiteen dat uitgeverijen ook bijdragen aan de symbolische productie, omdat ze hiërarchie aanbrengen in hun aanbod en zich inzetten om de boeken aantrekkelijker te maken voor de doelgroep. Daarmee zijn redacteurs medebepalend voor de wijze waarop een werk als literair wordt opgevat.

Tijdens het redactieproces heeft de redacteur tijd en sociale vaardigheden nodig om de auteur aandacht te geven, te inspireren en ervan te overtuigen dat aanpassingen het boek ten goede komen. Auteurs waarderen goede redactionele begeleiding en verlaten hun uitgeverij als ze ontevreden zijn, maar volgen ook regelmatig hun redacteur als ze van baan wisselen (Janssen, 2000, p. 75). Zo verlieten begin november 2017 ruim zestig jeugdboekenauteurs Querido Kind nadat de zeskoppige redactie was opgestapt (De Veen, 2017). Dit illustreert dat het de auteur minder om het uitgeefhuis zelf gaat en meer om de specifieke persoon met wie de auteur in contact staat.

Het vakmanschap en de creativiteit van de schrijver vragen een terughoudende wijze van redactie. Binnen de grenzen van fictie en non-fictie zijn er immers talloze opties om een effectieve tekst te schrijven. Redacteurs hebben de taak om de ideeën van de auteur zo goed mogelijk over het voetlicht te brengen (Coser et al., 1985, p. 122). Bij deze vorm van tekstevaluatie vormt de poëtische functie van taal het uitgangspunt. Redacteurs gaan ervan uit dat schrijvers met taal kunnen spelen via de dimensies klank, ritme, betekenis, zinsstructuur en opmaak (Brillenburg, Wurth, & Rigney, 2006, pp. 48-49).

Het is aan de redacteur om de juiste vragen te stellen of beargumenteerde verbetersuggesties te doen. Zo geeft Thomas Heerma van Voss aan dat er nauwelijks iets over is gebleven van de eerste versie van zijn tweede roman *Stern* (2013, uitgeverij Thomas Rap). Volgens de auteur heeft zijn redacteur Erik de Bruin de belangrijkste rol gespeeld in het herschrijfproces. Hij legde de schrijver een lijstje met negen vragen en opmerkingen op macroniveau voor, waarmee de auteur zelf aan de slag is gegaan (zie Figuur 1). De Bruin stelde zich terughoudend op: Heerma van Voss studeerde in die tijd zelf Nederlands en werkte zelf als redacteur.

- Doorwerking John: genoeg, helder? Niet te expliciet?,
- Niet teveel recapitulieren (die neiging is er te sterk, terugverwijzen naar eerdere scènes eigenlijk in alle delen boek).
- Genoeg Londense illusie breed gedragen vriendschap?
- Aanloop/rol Merel: niet teveel zwijgen, geen tekst te binnen schieten, etc.
- Postbode-scène misschien iets te karikaturaal?
- Tweede helft: iets te vaak benadrukt dat Bram een jochie is?
- Rol Merel in elke scène nog goed checken, niet teveel gewetensvrouw, maar daadwerkelijke aanwezigheid. (Ook bij adoptie geprobeerd iets aan te sterken.)
- Voornaamste punt dat ik nog niet helemaal getackeld heb: Merel stappen gezet, maar nog niet voldoende.
- John/Londen nog meer laten doorwerken bij interview? (Is vermoedelijk goed zo.)
- Heeft John, ondanks zijn 'leegte', wel genoeg gestalte? (Ik denk zelf: bijna.)

Figuur 1 Redactie-aanwijzingen manuscript Stern (Master Redacteur/editor, 2014, p. 14)

Idealiter werken zes mensen aan de tekst tijdens het proces van manuscript tot boek (Kuitert, 2008b, p. 19); in de praktijk zijn het er minimaal drie. De redacteur beoordeelt de inhoud, de structuur en de stijl van het manuscript en onderhoudt een goede werkrelatie met zijn auteur; hij werkt samen met de auteur totdat de tekst publicabel is. Nadat het volledige manuscript is overgedragen aan de bureauredactie, houdt de redacteur de supervisie. In de fase van de bureauredactie wordt het manuscript nagekeken op woordkeuze, zinsconstructie, spelling, interpunctie en huisstijl (Levy & Mole, 2017, p. 34), wordt de tekst persklaar gemaakt en de zetproef gecorrigeerd op fouten die er na de vormgeving in zijn geslopen. Deze drie taken waarbij de tekst wordt geperfectioneerd, worden in Nederland vaak uitbesteed aan freelancers en gecoördineerd door de bureauredactie.

Tijdens de hiervoor genoemde redactiefasen baseren redacteurs hun vragen en verbetervoorstellen op hun kennis van boeken, genres en publiek. Hoewel al in de inleiding duidelijk werd dat het aantal overgeleverde productiedossiers schaars is, kan met behulp van de in de inleiding genoemde reconstructies een impressie gegeven worden van de redactievoorstellen en -ingrepen van redacteurs bij Nederlandse uitgeverijen. De boekreconstructies uit het corpus bevestigen dat redacteurs zich terughoudend opstellen en rekening houden met de conventies van het tekstgenre en het effect van de afwijkingen ervan. Deze conventies qua tekstdoel, inhoud, opbouw, formulering en presentatie fungeren als schrijfmodellen voor auteurs en als leidraad voor de redacteur (Schuurs & De Groot, 2018, pp. 250-251). Wanneer auteurs hiervan afwijken gaat de redacteur na of deze afwijkingen

functioneel zijn. Zo beschrijft redacteur Jelte Nieuwenhuis (voorheen Atlas Contact) hoe auteur Lieke Marsman op zoek ging naar een nieuwe taal voor haar roman *Het tegenovergestelde van een mens* (2017):

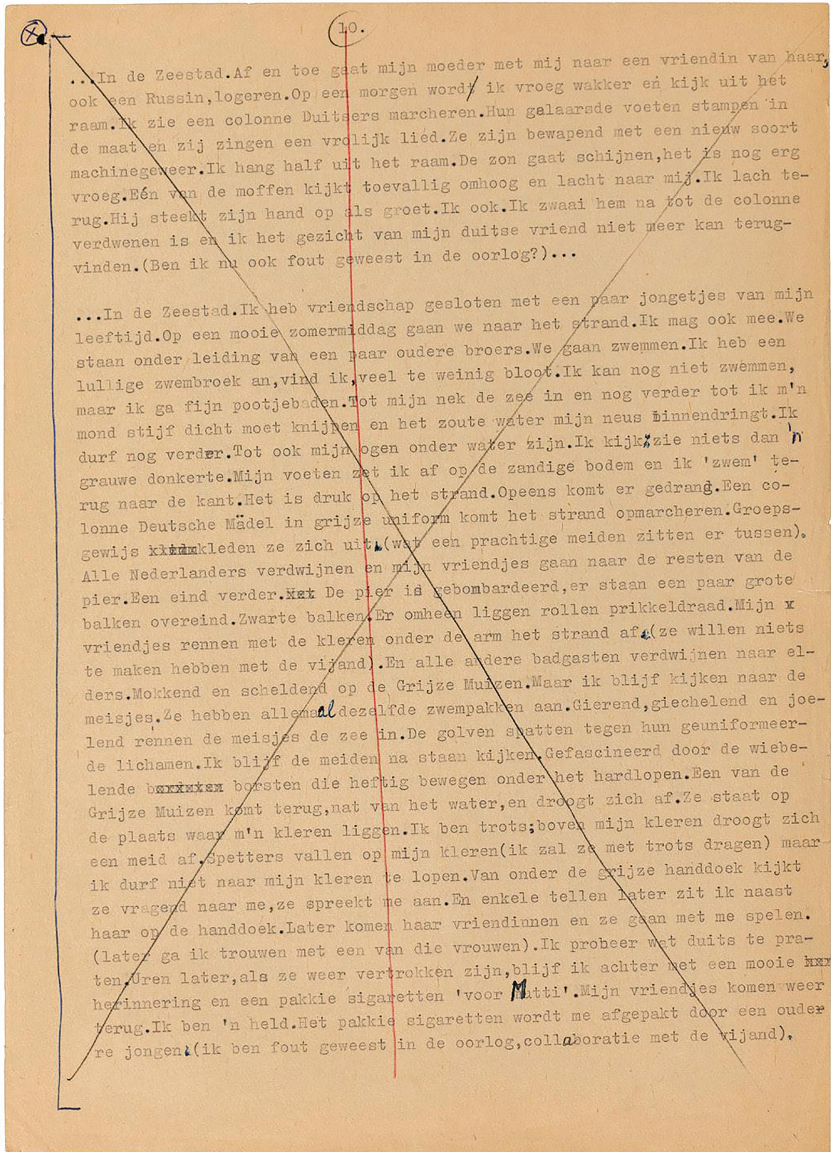
[H]oe vind je een nieuwe taal? Een traditionele roman is vaak al ingewikkeld genoeg. In dat geval spreken redacteurs vaak over het vinden van de juiste toon. Dat is de ruwe grondstof waarmee schrijvers en redacteurs werken. Maar Lieke zorgt voor een heel nieuw genre. Een mix van proza, poëzie en essayistiek. Dat klinkt kunstmatiger dan het is, want in de poëzie waarmee ze doorbrak zaten al allerlei narratieve en beschouwende elementen. (Nieuwenhuis, 2017)

Aan de hand van de drie hoofdvragen uit het CCC-model (Renkema, 2002, pp. 178-179; Renkema, 2012, pp. 28-29) wordt hierna aan de hand van enkele casussen uit het corpus geïllustreerd hoe de redactionele keuzes in deze casussen tijdens de manuscript- en kopijbeoordeling te verklaren zijn:

- 1 Is de afstemming tussen schrijver en lezer adequaat?
- 2 Blijft de schrijver bij zijn keuzes?
- 3 Worden de algemene regels voor het taalgebruik gevolgd: worden de genrerregels goed toegepast, kloppen de feiten, zijn de zinnen correct, enzovoort?

Op basis van de genreconventies en de functionaliteit van eventuele afwijkingen ervan geeft de redacteur feedback. Zo worden onder het eerste criterium de redactieaanwijzingen geschaard die betrekking hebben op de afstemming van de tekst op het doel en de doelgroep. Hierbij gaan redacteurs bijvoorbeeld na of de inhoud van de tekst voldoende is afgestemd op de lezer en – in dit specifieke geval – aansluit bij het imago van de uitgeverij. Zo is in de veilingcatalogus *De onverbiddelijke veiling van het manuscript van Ik Jan Cremer* (Blom, 2010) te zien dat de redacteur meerdere passages heeft geschrapt die als ‘fout’ of ‘antisemitisch’ kunnen worden bestempeld. Ze passen niet in de tijd waarin het boek verscheen – kort na de Tweede Wereldoorlog – en niet bij een uitgeverij die is voortgekomen uit het verzet. Het heeft redacteur Oscar Timmers en uitgever Geert Lubberhuizen van De Bezige Bij begin jaren 60 veel inspanning gekost voordat Jan Cremer instemde met de voor de uitgeverij noodzakelijke redactie-ingrepen. Zo is bijvoorbeeld pagina 10 (zie Figuur 2) geschrapt. De redacteur heeft een anekdote gecensureerd waarin de jonge Jan naar het strand gaat en heult met een:

'colonne Deutsche Mädels' en bij één van hen op een handdoek komt zitten. Hij neemt een pakje sigaretten aan 'voor Mutti'. Laatste regels: 'Ik ben 'n held. Het pakje sigaretten wordt me afgepakt door een oudere jongen. (ik ben fout geweest in de oorlog, collaboratie met de vijand.)' (Blom, 2010, p. 27)



Figuur 2 Verwijderde pagina uit het typescript van Ik Jan Cremer (Co4242; Blom, 2010, p. 27)

Ook in andere boekreconstructies worden uiteenlopende redactie-aanwijzingen op correspondentieniveau gesignaleerd. Kluuns debuut *Komt een vrouw bij de dokter* is na elf versies gepubliceerd. Op voorspraak van zijn redacteur is bij de eerste redactierondes het aantal personages verminderd, en zijn de compositie van de roman, de stijl en de woordkeuze aangepast. De auteur geeft zelf aan dat er veel tekst is gesneuveld om deze beter af te stemmen op het doel en het publiek.

Ruim 400 dichtbeschreven A4-tjes met hopeloos veel totaal overbodige en bovendien slecht uitgewerkte personages (het leek met al die namen die je moest onthouden af en toe wel een familieroman van Couperus, maar dan – gelukkig – veel leuker en – helaas – minder goed geschreven), ellenlange beschrijvingen van ontelbare ziekenhuisbezoeken, breed uitgemeten puberaal beschreven seksscènes en vooral veel onbegrijpelijk debutantengefilosofer over de zin van het leven. (Kluun, 2009)

Van een andere aard zijn de redactie-aanwijzingen aan Frits van Oostrom die externe redacteur Aggie Langedijk heeft opgenomen in de derde versie van *Stemmen op schrift* (2005, uitgeverij Prometheus). In Figuur 3 is te zien hoe zij suggesties doet die het tekstbegrip ten goede moeten komen; met

hooftse hofstijl inrijdels (honderd jaar/een eeuw na Veldeke) tot f huisstijl was geworden, ook al hebben wij daarvan nauwelijks een schijn van documentatie, anders dan behalve door een buitenkans als deze.

Wél zelfstandig overgeleverd – maar vraag niet hoe! – heerlijk gehavend namelijk – zijn de *Lundse liederen*, zscstien bewaard in fragmenten in het Zweedse Lund, en uit een ander handschrift van dezelfde ‘bundel’ nog eens acht stuks te Brussel. Sinds hun ontdekking in 1926 zijn ze erkend als heel bijzondere poëzie, nagenoeg zeker van een en dezelfde dichter, die zich lijkt te hebben toegelegd op steeds dezelfde vorm als handelsmerk. Alle zijn gedichten vertonen een driedeling, met in de kern een strofe waarin een lyrisch ik, zoals bij Veldeke, zijn persoonlijke minnesentiment uitdrukt spreekt. – maar hier omkransd door openings- en slotstrofen het is een driedeling, met in het midden een lyrisch ik, dus dan moeten die andere strofen wel begin- en eindstrofe zijn? die in verwante variaties die stemming tillen op een algemeen-menselijk plan, door met generaliserende uitspraken over de liefde. De dichter heeft hiermee een zeer vernuftige vorm gevonden voor de wisselwerking tussen individuele en collectieve liefdeservaring, wat van meet af aan het spanningsveld was waarin zich het hooftse lied afspeelde: dit lijkt me geen spanningsveld, want het publiek herkent zich in het minnesentiment van het ik?) de wisselwerking tussen de (quasi-

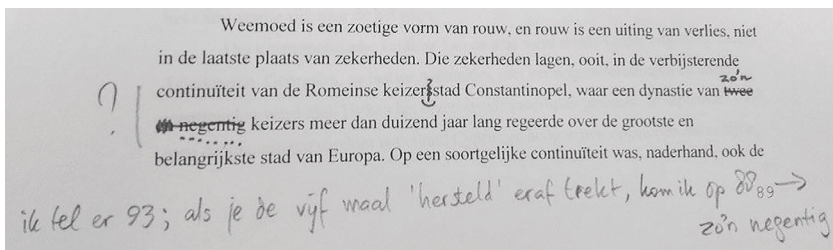
Figuur 3 Fragment met redactie-aanwijzingen vierde versie *Stemmen op schrift*

potlood heeft Van Oostrom aangegeven welke hij gaat overnemen en welke niet (Van Oostrom, z.d.).

Het tweede criterium uit het CCC-model betreft de consistentie van de tekst. Eenmaal gemaakte keuzes moeten consequent worden doorgevoerd, zodat het leesgemak en tekstbegrip worden bevorderd. In Vromans bundel *Huis en tuin. Fabels en strips* (1979, uitgeverij Querido) worden fabels en strips gecombineerd. Deze zijn eerder in *De Revisor* en *Hollands Maandblad* verschenen en worden voor deze bundel aangevuld met nieuw materiaal. Vroman en Querido-uitgever Reinold Kuipers spreken af 'de fabels en tekeningen vanzelf te laten groeien tot [Kuipers] het totaal voldoende vindt voor een boekje' (Vroman, 1978). Vroman stelt voor om de fabels en strips te combineren. Kuipers vindt dit geen goed idee, omdat de illustraties niet passen bij de gedichten waar ze bij staan: deze inconsistentie kan verwarrend zijn voor de lezer. Hij wil de strips daarom combineren tot een afdeling in het midden van het boek. Vroman gaat akkoord onder de voorwaarde dat de strips achterin het boek worden gecombineerd.

Tijdens het schrijven van *Hitler. Opkomst en ondergang van een Duits politicus* (2012, uitgegeven bij Prometheus) heeft redacteur Mascha Holman bij historicus Willem Melching een verbeterpunt aangekaart dat de consistentie van de formulering betreft. Ze adviseerde Melching om Hitler en Adolf niet door elkaar te gebruiken.

Ook het derde criterium van het CCC-model – de correcte toepassing van de algemene regels voor het taalgebruik (inclusief genre- en taalconventies), met eventuele beredeneerde afwijkingen ervan, en ook de weergave van feiten – kan geïllustreerd worden met een voorbeeld. In het manuscript van Boekenweekgeschenk *De Brug* (2006) zijn twee correcties te zien die aansluiten bij het beoordelingscriterium correctheid. De eerste correctie betreft een fout die Renkema (2012, p. 29) schaaft onder het kopje 'onjuiste weergave van feiten'. In Figuur 4 is te zien dat dat het genoemde aantal keizers dat over Constantinopel regeerde, inhoudelijk is gecorrigeerd, net als de spelfout in keizerstad. Geert Mak heeft beide verbetervoorstellen geaccepteerd.



Figuur 4 Correctie manuscript *De Brug* (Productiedossier A 4765 [NG], 2006; Master Redacteur/editor, 2017)

Conclusie en discussie

In dit exploratieve onderzoek is gekeken naar criteria die redacteuren hanteren bij de acquisitie en redactie van boeken bij Nederlandse algemene uitgeverijen. In de acquisitiefase gaat de redacteur op basis van zijn kennis van teksten en belezenheid na welke teksten in aanmerking komen voor uitgave. Uit literatuuronderzoek blijkt dat de commerciële potentie van het werk en de auteur, de aansluiting bij het fonds en de traditie van de uitgeverij, de gemaakte afspraken met de auteur en de investering in redactie en redigeerbaarheid van de auteur de criteria vormen voor tekstbeoordeling.

De redacteur hanteert ook criteria bij tekstredactie en draagt bij aan de optimalisering van de tekstkwaliteit binnen de mogelijkheden die hij heeft. Hij stelt zich niet alleen terughoudend op, maar maakt ook strategische afwegingen op basis van zijn kennis van de auteur en zijn kennis van teksten. De voorbeelden die gerelateerd zijn aan de correspondentie, consistentie- en correctheidsvragen uit het CCC-model (Renkema, 2002; 2012) geven de indruk dat de redacteur beredeneerde wijzigingen voorstelt die te relateren zijn aan tekstkwaliteit, en afwijkingen van genreconventies intact laat als ze aansluiten bij het doel en de doelgroep, en consequent zijn doorgezet. Dik Zweekhorst (redacteur Em. Querido's Uitgeverij) bevestigt dit: 'Dat geldt voor elk boek dat je redigeert. Als redacteur zie je al aan de eerste paar bladzijdes welk register de auteur hanteert, en dit moet je consequent doorvoeren.'

Deze exploratieve studie geeft een eerste indruk van de bijdrage van de redacteur als tekstbeoordelaar binnen de Nederlandse algemene uitgeverij. Om de teksttransformatie tijdens het redactieproces en de inbreng van de redacteur nauwkeuriger te onderzoeken en vanuit taalbeheersingsperspectief te verklaren is binnen het Huygens ING per 1 september 2018 het onderzoeksproject *Track changes* gestart waarbinnen op empirische wijze redactiepatronen worden onderzocht en verklaard in moderne Nederlandstalige fictie en literaire non-fictie.

Noot

1. Deze reconstructies zijn tot stand gekomen in samenwerking met studenten van de masteropleiding Redacteur/editor (lichting 2013-2014, 2014-2015, 2015-2016, 2016-2017, 2017-2018) aan de Universiteit van Amsterdam.

Bibliografie

- Aalbrecht, H., & Wagenaar, P. (2015). *Handboek voor de redacteur*. Amsterdam: Boom.
- Becker, H. S. (2008). *Art Worlds*. Berkely/Los Angeles/Londen: University of California.
- Blom, O. (2010). *De onverbidelijke veiling van het manuscript van Ik Jan Cremer. (veilingcatalogus)*. Amsterdam: Adams Amsterdam Auctions.
- Brillenburg Wurth, K., & Rigney, A. (Red.) (1985). *Het leven van teksten. Een inleiding tot de literatuurwetenschap*. Amsterdam: Amsterdam University Press.
- Buschenhenke, F. (z.d.). *Vragenlijst Het Literaire Werk 2.0: De resultaten*. Geraadpleegd op 30 juni 2018, <https://www.huygens.knaw.nl/het-literaire-werk-2-0/>
- Coser, L. A. (1975). Publishers as gatekeepers of ideas. *The Annals of the American Academy of Political and Social Science*, 421(1), 14-22.
- Coser, L. A., Kadushin, C., & Powell, W. W. (1985). *The Culture and Commerce of Publishing*. Chicago: University of Chicago Press.
- Cremer, J. (1963). *Ik Jan Cremer (typoscript, Co4242)*. Den Haag, Literatuurmuseum.
- Devitt, A. J. (2000). Integrating rhetorical and literary theory of genres. *College English*, 62(6), 696-718.
- Dorrestein, R. (2018). *Dagelijks werk. Een schrijversleven*. Amsterdam: Uitgeverij Podium.
- Fortuin, A. (2015). *Geert van Oorschoot, uitgever*. Amsterdam: Van Oorschoot.
- Franssen, T. (2012). Leren lezen: Over de beoordeling van manuscripten door acquirierend redactoren. In K. Absillis, & K. Humbeek (Red.), *Literatuurwetenschap en uitgeverijonderzoek* (Cahier voor literatuurwetenschap, nr. 4) (pp. 61-77). Gent: Ginkgo Academia Press.
- Franssen, T. (2015). *How Books Travel. Translation Flows and Practices of Dutch Acquiring Editors and New York Literary Scouts, 1980-2009*. Proefschrift, Universiteit van Amsterdam.
- Glas, F. de (2006). The role of the editor in the fiction publishing branch: Towards the institutional analysis of a profession. *Frame: Tijdschrift voor Literatuurwetenschap*, 18(3), 7-25.
- Janssen, S. (2000). Onderzoek naar twintigste-eeuwse uitgeverijen: Een stand van zaken. *Jaarboek voor Nederlandse Boekgeschiedenis*, 7, 65-79.
- Kluun (z.d.). *The Directors Cut*. Geraadpleegd op 25 augustus 2009, <http://kluun.nl/>
- Kuitert, L. (2008a). De uitgeverij en de symbolische productie van literatuur: Een historische schets 1800-2008. *STILET*, 10(2), 67-87.
- Kuitert, L. (2008b). *Over redactie*. Amsterdam: Augustus.
- KVB Boekwerk (z.d.). *Boekwerk Monitor 2017*. Geraadpleegd op 28 februari 2018, <https://www.kvb-boekwerk.nl/>
- Levy, M., & Mole, T. (2017). *The Broadview Introduction in Book History*. Peterborough, Ontario: Broadview Press.
- Lewin, K. (1951). *Field Theory in the Social Sciences*. New York: Harper and Row.
- Master Redacteur/editor (2014). *Achter de schermen, op de voorgrond*. Amsterdam: Universiteit van Amsterdam.
- Master Redacteur/editor (2017). *Binnenstebuiten: Het literaire redactieproces onthuld*. Amsterdam: Universiteit van Amsterdam.
- Master Redacteur/editor (2018). *Hemelsbrede harmonie: De verbindende rol van het CCC-model in de redactie van verschillende genres*. Amsterdam: Universiteit van Amsterdam.
- Nieuwenhuis, J. (z.d.). *Een nieuwe taal*. Geraadpleegd op 23 november 2017, <https://boekblad.nl/>
- Oostrom, F. van (z.d.). *The making of* Geraadpleegd op 10 januari 2018, <http://www.stemmenopschrift.nl/themakingof>
- Poel, K. Van de, Carstens, W. A. M., & Linnegar, J. (2012). *Text Editing. A Handbook for Students and Practitioners*. Brussel: University Press Antwerp.

- Productiedossier A 4765 [NG]*. Den Haag, Literatuurmuseum. Bronnenmateriaal. Niet nader gecatalogiseerde archivalia Geert Mak.
- Renkema, J. (2002). Over smaak valt goed te twisten: Een evaluatiemodel voor tekstkwaliteit. In F. van Eemeren, P. van den Hoven, C. Jansen, & P. J. Schellens (Red.), *Tussenstand: 25 jaar Tijdschrift voor Taalbeheersing* (pp. 177-190). Assen: Van Gorcum. (Eerder verschenen in *Tijdschrift voor Taalbeheersing*, 18 (1996), 314-338.)
- Renkema, J. (2012). *Schrijfwijzer* (5e dr.). Amsterdam: Boom.
- Schuurs, U., & Groot, E. de (2018). Genre-analyse. In J. Karreman & R. van Enschoot (Red.), *Tekstanalyse: Methoden en toepassingen* (pp. 239-293). Assen: Van Gorcum.
- Sharpe, L. T., & Gunther, I. (1994). *Editing Fact and Fiction. A Concise Guide to Book Editing*. Cambridge: Cambridge University Press.
- Veen, T. de (2017). Redactie Querido Kinderboeken begint nieuwe uitgeverij. *NRC Handelsblad*, 8 november 2017.
- Vroman, L. (1978). *Brief Leo Vroman aan Reinold Kuipers 29 mei 1978*. (correspondentie, 19 VRO). Den Haag: Literatuurmuseum.

Een pragma-dialectische reconstructie van de discussiebijdragen van arts en patiënt in ‘shared decision-making’

A. Francisca Snoeck Henkemans en Jean H.M. Wagemans
Universiteit van Amsterdam

TVT 41 (1): 281–293

DOI: 10.5117/TVT2019.1.019.SNOE

Abstract

A pragma-dialectical reconstruction of contributions to the discussion by both doctor and patient in ‘shared decision-making’

Medical consultation nowadays generally proceeds in accordance with the process of shared decision-making (SDM). In the literature on medical communication, the general characteristics of this process seem to be agreed upon. With respect to the specific division of labour between doctor and patient, however, different alternatives are suggested. In this paper we approach the process of SDM from the perspective of the pragma-dialectical theory of argumentation. We translate the contributions of doctor and patient to this process in terms of the ideal model of a critical discussion, taking into account both the general characteristics and the different alternatives for the division of labour proposed in the literature. The resulting analytical tool enables a more precise analysis of the process of SDM and creates a starting point for a well-founded evaluation of the contributions of doctor and patient to this process.

Keywords: argumentation, doctor-patient communication, medical consultation, pragma-dialectics, shared decision-making, health communication

Inleiding

Voor gesprekken tussen arts en patiënt wordt tegenwoordig vrij algemeen het model van ‘shared decision-making’ (SDM) als ideaal gezien (Stiggelbout, Pieterse, & De Haes, 2015). Volgens dit model nemen arts en patiënt in gezamenlijk overleg een beslissing over wat de beste behandeling voor de patiënt is. Dit in tegenstelling tot oudere modellen voor arts-patiëntcommunicatie zoals het ‘paternalistische model’, waarin de arts beslist wat de beste behandeling voor de patiënt is, of het ‘informed-decision making model’, waarin de besluitvorming over de behandeling volledig aan de patiënt wordt overgelaten (Charles, Gafni, & Whelan, 1997, pp. 682-683).

Volgens de klassiek geworden beschrijving van Charles, Gafni en Whelan (1999, p. 652) heeft SDM in ieder geval als kenmerken dat arts en patiënt bij de besluitvorming over de behandeling betrokken zijn, informatie met elkaar delen, behandelvoorkeuren naar voren brengen, en gezamenlijk tot overeenstemming komen over de uiteindelijke keuze voor een behandeling.

In de medische literatuur kan men zich doorgaans vinden in bovenstaande algemene karakterisering van het proces van SDM. Tegelijkertijd zijn er uiteenlopende voorstellen gedaan voor de concrete invulling van de taakverdeling tussen arts en patiënt in de achtereenvolgende fasen van het besluitvormingsproces. Als gevolg van deze verscheidenheid aan opvattingen bestaat er in de literatuur weinig overeenstemming over de beoordeling van de kwaliteit van medische consultaties.

In dit artikel analyseren we het proces van SDM vanuit het perspectief van de pragma-dialectische argumentatietheorie (Van Eemeren & Grootendorst, 1984, 2004). Daardoor wordt het mogelijk de verschillende bijdragen van arts en patiënt aan dat proces op een nauwkeurige manier te reconstrueren. Zo’n reconstructie kan vervolgens als uitgangspunt dienen voor een wetenschappelijk verantwoorde beoordeling van de kwaliteit van die bijdragen.

De opzet van dit artikel is als volgt. In de volgende paragraaf beschrijven we de algemene uitgangspunten en kenmerken van SDM. Vervolgens geven we een overzicht van de verschillende opvattingen over de precieze taakverdeling tussen arts en patiënt. Op basis van deze inventarisatie van de algemene en specifieke kenmerken van SDM vertalen we in de daaropvolgende paragraaf het besluitvormingsproces in termen van het pragma-dialectische ideaalmodel van een kritische discussie. In de laatste paragraaf reflecteren we op de bruikbaarheid van het ontwikkelde analyse-instrument voor het reconstrueren en beoordelen van de bijdragen van de arts en de patiënt aan het proces van SDM.

Algemene uitgangspunten en kenmerken van SDM

Het gespreksmodel van SDM is gebaseerd op de gedachte dat aan de patiënt een grote mate van autonomie moet worden toegekend (Charles et al., 1997). In dit opzicht verschilt SDM van het vroeger gangbare paternalistische gespreksmodel, dat ervan uitgaat dat de patiënt zich volledig zal verlaten op de professionaliteit van de arts (zie bijvoorbeeld Emanuel & Emanuel, 1992, p. 2221; Kaba & Sooriakumaran, 2007).

Het ideaal van autonomie van de patiënt is in het model van SDM tot uitdrukking gebracht doordat de patiënt en niet de arts het recht is gegeven om een beslissing te nemen over de behandeling.¹ Omdat het verschil dat daardoor tussen de deelnemers wordt gecreëerd betrekking heeft op een ongelijkheid in de rechten of plichten die aan hen zijn toebedeeld, kan in dit verband gesproken worden van een 'deontische' asymmetrie (zie bijvoorbeeld Lindström & Weatherall, 2015; Stevanovic, 2013).

Een ander gegeven dat de vormgeving van het model van SDM heeft beïnvloed is het niveauverschil in biomedische kennis dat in het algemeen tussen arts en patiënt bestaat. De grootte van dit niveauverschil, dat ook wel als een 'epistemische' asymmetrie wordt aangemerkt (zie bijvoorbeeld Raymond, 2014), verschilt uiteraard van patiënt tot patiënt. De laatste jaren is het gemiddelde niveauverschil wellicht iets verkleind door de opkomst van websites met medische informatie, die veelal voorafgaand aan het gesprek door de patiënt worden geraadpleegd (Van Klaveren, 2017).

Terwijl dus verondersteld mag worden dat de arts het meeste biomedische kennis heeft, is de situatie wat de kennis van de voorkeuren van de patiënt betreft precies omgekeerd. Het is immers de patiënt, en niet de arts, die de meeste kennis heeft over zijn persoonlijke voorkeuren voor bepaalde behandelwijzen en zijn ervaringen met de effecten ervan (Raymond, 2014, p. 39). Dit niveauverschil kan als een tweede epistemische asymmetrie worden aangemerkt.

Ten slotte is het voor de invulling van de taakverdeling tussen arts en patiënt van belang te onderkennen dat het proces van SDM plaatsvindt in de context van het medisch consult. Het initiatief tot en de verantwoordelijkheid voor het uitvoeren van de verschillende stappen in dat proces zal daardoor in de meeste gevallen bij de arts liggen.

In de medische literatuur bestaan uiteenlopende opvattingen over de aard, het aantal en de volgorde van de stappen waaruit het proces van SDM bestaat. In de onderstaande globale beschrijving van dat proces volgen we het zogenoemde 'geïntegreerde model van SDM', dat tot stand is gekomen op basis van een systematisch onderzoek naar conceptuele definities van

SDM in een representatief aantal artikelen in de tot dan toe verschenen medische literatuur.

Het geïntegreerde model van SDM bestaat uit verschillende elementen of procedurele stappen, waarvan sommige als ‘essentieel’ en andere als ‘ideaal’ worden bestempeld (Makoul & Clayman, 2006, pp. 305-306). Tot de essentiële elementen behoren (1) het definiëren of uitleggen van het probleem; (2) het presenteren van de opties; (3) het bediscussiëren van de voor- en nadelen van die opties; (4) het bespreken van de waarden en voorkeuren van de patiënt; (5) het bespreken van de zelfstandigheid en capaciteiten van de patiënt; (6) het naar voren brengen van medische kennis en aanbevelingen; (7) het controleren of de patiënt het begrepen heeft; (8) het nemen of expliciet uitstellen van de beslissing; (9) het maken van een vervolgspraak.

Uit deze beschrijving is duidelijk dat het geïntegreerde model van SDM niet vastlegt wat de precieze taakverdeling is tussen arts en patiënt. Volgens Makoul en Clayman kan de mate waarin de verschillende stappen gedeeld worden als een continuüm worden opgevat: ‘It may be helpful to envision the degree of sharing as a continuum with physicians leading the discussion and making decisions at one end, patients leading the discussion and making decisions at the other, and truly shared discussion and decision-making in the middle’ (2006, pp. 306-307). Om meer zicht te krijgen op de verscheidenheid aan opvattingen over de taakverdeling tussen arts en patiënt, geven we in de volgende paragraaf een overzicht van de belangrijkste invullingsvarianten die in de medische literatuur naar voren zijn gebracht.

De taakverdeling tussen arts en patiënt

Hoewel het delen van voorkeuren door zowel patiënt als arts in de oorspronkelijke beschrijving van SDM van Charles et al. (1997) als essentieel wordt gezien, wijzen sommige auteurs op het mogelijke gevaar dat de arts te sturend optreedt en zijn eigen voorkeuren te veel laat doorschemeren in de presentatie van de behandelopties. Een dergelijke manier van handelen zou tot een verstoring van het proces van SDM leiden (Engelhardt et al., 2016, p. 56).

Andere auteurs zijn van mening dat artsen wel degelijk het recht hebben om hun mening kenbaar te maken omdat er anders niet echt sprake is van *shared* decision-making (Emanuel & Emanuel, 1992). Juist het feit dat de arts zijn mening geeft en deze niet achterhoudt, zou kenmerkend zijn voor het collaboratieve karakter van SDM.

Een belangrijke factor die bepaalt of de arts zijn eigen voorkeuren mag verdedigen tegenover de patiënt of zich juist op de vlakte moet houden, lijkt de aard van de beslissing te zijn. In gevallen waarin er medisch gezien een duidelijke voorkeur is voor een bepaald type behandeling, zou volgens sommige auteurs het beïnvloeden van patiënten zelfs geboden zijn (Dubov, 2015, p. 501). Daarentegen kan het ook voorkomen dat er uit een aantal medisch gezien gelijkwaardige behandelingen gekozen moet worden, een situatie waarin volgens Engelhardt et al. (2016, p. 56) het nemen van een gezamenlijke beslissing bij uitstek op zijn plaats is. In zo'n geval zouden de wensen van de patiënt de doorslag moeten geven, en dient de arts de patiënt niet in de richting van de keuze voor een bepaalde behandeling te sturen (tenzij de patiënt hem uitdrukkelijk verzoekt om de beslissing te nemen).

Een andere kwestie waarover in de literatuur over SDM verschillende opvattingen bestaan, is de vraag of de arts de inspannings- of zelfs de resultaatverplichting heeft om de preferenties van de patiënt boven tafel te krijgen en in hoeverre de patiënt zelf verplicht is deze preferenties kenbaar te maken. In Nederland is de patiënt volgens de Wet op de Geneeskundige BehandelingsOvereenkomst (WGBO) verplicht om de zorgaanbieder correct en zo volledig mogelijk te informeren (BW 7, artikel 452). Om wat voor soort informatie het hier gaat wordt verder niet gespecificeerd, maar in de context van SDM zou het – behalve om informatie van medische aard – daarbij in ieder geval ook om de doelen, waarden en voorkeuren van de patiënt gaan. Volgens Ishikawa, Hashimoto en Kiuchi (2013, p. 149) is het de taak van de patiënt om zijn eigen waarden en voorkeuren kenbaar te maken zodat deze bij de beoordeling van de door de arts aangedragen medische feiten in beschouwing kunnen worden genomen. Tegelijkertijd vinden zij dat het initiatief voor het boven tafel krijgen van die preferenties in het proces van SDM bij de arts moet worden gelegd.

Een laatste punt waarover in de literatuur verschillende zienswijzen bestaan is de vraag of de taakverdeling tussen arts en patiënt – en dan met name de wijze waarop de beslissing genomen zal worden – al dan niet vooraf expliciet afgesproken dient te worden. Towle en Godolphin (1999, p. 767) zien het als een essentiële taak voor de arts om vooraf vast te stellen welke rol de patiënt wenst te hebben bij het nemen van een beslissing over wat de beste behandeling is. Andere auteurs, waaronder Charles et al. (1999, p. 656), menen dat hier ook op een impliciete manier tijdens het gesprek overeenstemming over kan worden bereikt. Stiggelbout et al. (2015, p. 1174) vinden het vooral van belang dat de arts vooraf expliciet aan de patiënt

duidelijk maakt dat er een echte keuze te maken is en dat de voorkeuren van de patiënt daar een belangrijke rol in zullen spelen.

Een pragma-dialectische reconstructie van het besluitvormingsproces

Nu duidelijk is welke opvattingen er in de medische literatuur bestaan over de algemene karakteristieken van het proces van SDM en welke voorstellen daarin zijn gedaan voor de taakverdeling tussen arts en patiënt, gaan we over tot het vertalen van het besluitvormingsproces in termen van het pragma-dialectische 'ideaalmodel van een kritische discussie' (Van Eemeren & Grootendorst, 1984, 2004). Voor elk van de vier discussiefasen die in dit model worden onderscheiden geven we een reconstructie van de mogelijke discussiebijdragen van de arts en de patiënt in het proces van SDM. We houden daarbij ook rekening met de eerdergenoemde invullingsvarianten van de taken die de discussianten in dit proces geacht worden te vervullen.

Confrontatiefase

De eerste discussiefase die in het ideaalmodel wordt onderscheiden is de 'confrontatiefase'. In deze fase moet worden vastgesteld of er sprake is van een verschil van mening en wat de aard daarvan is (Van Eemeren & Grootendorst, 2004, pp. 60, 135). Van een verschil van mening is al sprake wanneer er niet van uitgegaan kan worden dat een standpunt van één van de partijen bij voorbaat aanvaardbaar zal zijn voor de andere partij. Op basis van de hierboven besproken algemene kenmerken en varianten van het proces van SDM zijn de volgende typen verschillen van mening mogelijk tussen arts en patiënt.

In alle opvattingen over SDM dient de arts de standpunten te verdedigen dat de verschillende behandelingsmogelijkheden (B_1 - B_n) elk voor zich vanuit medisch gezichtspunt een acceptabele keuze zijn. Aangezien aangenomen mag worden dat de patiënt ten opzichte van elk van deze standpunten nog geen positie heeft ingenomen, is er sprake van een niet-gemengd meervoudig verschil van mening. De noodzaak tot het verdedigen van deze standpunten vloeit voort uit de verplichting die de arts heeft om de patiënt te laten weten uit welke medisch acceptabele behandelingen deze kan kiezen. De arts heeft dan de bewijslast voor het feit dat voor elk van de genoemde behandelingen de negatieve effecten medisch gezien opwegen tegen de positieve effecten.

Volgens sommige opvattingen van SDM zou de arts geen standpunt in moeten nemen over welke van de genoemde behandelingen het meest geschikt is, maar dient hij deze keuze over te laten aan de patiënt. Volgens andere opvattingen kan de arts echter wel aangeven welke behandeling in zijn ogen de voorkeur verdient. In een situatie waarin de patiënt hier nog geen mening over heeft, is het verschil van mening dat daardoor ontstaat te karakteriseren als niet-gemengd enkelvoudig. De arts brengt in dit geval het standpunt naar voren dat één van de behandelingsmogelijkheden medisch gezien de beste keuze is. Dat dit zo is, zou moeten worden verdedigd op grond van het feit dat volgens de arts de afweging van negatieve en positieve effecten bij die mogelijkheid het gunstigst uitpakt. Zoals uit de bespreking van de literatuur over SDM bleek, zou de arts een dergelijk standpunt in principe alleen naar voren mogen brengen wanneer er medisch gezien geen gelijkwaardige behandelingsmogelijkheden zijn.² Wanneer de patiënt, nadat de arts een voorkeur voor een bepaalde behandeling heeft uitgesproken, laat weten het niet met deze keuze eens te zijn, verandert het verschil van mening van niet-gemengd enkelvoudig naar gemengd enkelvoudig. De arts neemt dan een positief standpunt in met betrekking tot een bepaalde behandelingskeuze, en de patiënt neemt hiertegenover een negatief standpunt in. Als de patiënt alleen twijfel naar voren brengt met betrekking tot het standpunt van de arts, is het verschil van mening opnieuw niet-gemengd enkelvoudig. De patiënt heeft bij twijfel zelf geen standpunt ingenomen en dus ook geen verdedigingsplicht.

Op grond van de kenschets van SDM moet aangenomen worden dat het verschil van mening in veel gevallen in de loop van het beslissingsproces verandert. Dit gebeurt in die gevallen waarin de arts aanvankelijk geen voorkeur uitspreekt, maar alleen aangeeft dat er een aantal medisch gelijkwaardige behandelingsmogelijkheden zijn. Na de voorkeuren van de patiënt gehoord te hebben, kan de arts dan het standpunt verdedigen dat een bepaalde behandeling, gezien de voorkeuren van de patiënt, de beste keuze is. Op die manier kan de arts aan de discussie deelnemen zonder de patiënt al bij voorbaat in een bepaalde richting te sturen ('Nu ik hoor wat voor u belangrijk is, denk ik dat optie B_i de beste keuze voor u is'). Het verschil van mening verandert dan van niet-gemengd meervoudig naar niet-gemengd enkelvoudig. Ook wanneer de patiënt zelf een standpunt inneemt over wat de beste behandeling is ('Gezien wat voor mij belangrijk is, denk ik dat optie B_i de beste keuze is'), ontstaat een niet-gemengd enkelvoudig verschil van mening. Ten slotte kan het geschil ook gemengd enkelvoudig worden. Dit gebeurt wanneer arts en patiënt van mening verschillen over wat, gezien de voorkeuren van de patiënt, de beste optie is.

Openingsfase

In de openingsfase van de discussie moeten afspraken gemaakt worden over de uitgangspunten voor de discussie en over de procedure (Van Eemeren & Grootendorst, 2004, pp. 60, 137). Een belangrijke procedurele stap is het vaststellen van de bewijslast. De verdeling van de bewijslast hangt samen met het type verschil van mening. In een niet-gemengd verschil van mening heeft maar een van de partijen de bewijslast, namelijk de partij die een standpunt naar voren heeft gebracht. In een gemengd verschil van mening hebben beide partijen de bewijslast voor hun eigen standpunt.

Bij de genoemde typen verschillen van mening die tijdens SDM kunnen ontstaan, heeft de arts bij het eerste type geschil de bewijslast voor het feit dat elk van de opties medisch gezien een acceptabele keuze is. Bij het tweede type geschil, dat ontstaat wanneer er geen gelijkwaardige alternatieven zijn, heeft de arts de bewijslast voor het feit dat zijn keuze medisch gezien de beste is. Bij het derde type geschil, waarbij hetzij de arts, hetzij de patiënt het standpunt inneemt dat een van de behandelkeuzes de beste keuze is gezien de voorkeuren en waarden van de patiënt, ligt het iets genuanceerder. In die situatie geldt voor de arts dat hij dan ook de bewijslast voor zijn standpunt heeft, maar voor de patiënt niet per se. Omdat de patiënt uiteindelijk het recht heeft om te beslissen voor welke van de medisch verantwoorde behandelingen hij kiest, is hij niet verplicht zijn standpunt te onderbouwen. Als hij dit niet doet, is de consequentie wel dat het verschil van mening niet wordt opgelost, maar alleen beslecht.

In sommige beschrijvingen van SDM wordt opgemerkt dat de arts, vanwege zijn rol als gespreksleider, ervoor moet zorgen dat de patiënt zijn voorkeuren kenbaar maakt. In meer recente beschrijvingen wordt de rol van de patiënt hierin benadrukt. Argumentatief gesproken maakt het geen verschil welke procedure gevolgd wordt bij het inbrengen van de voorkeuren van de patiënt in de discussie.

Een ander belangrijk procedureel uitgangspunt voor SDM is dat in de beoordeling van de behandelkeuzes als beslisregel moet worden gehanteerd dat van de medisch verantwoorde behandelingen die behandeling de voorkeur heeft die het meest recht doet aan de voorkeuren en waarden van de patiënt.

Tot de procedurele afspraken kunnen ook de in de literatuur over SDM genoemde afspraken over hoe het besluitvormingsproces als geheel zal verlopen gerekend worden. In de literatuur verschilt men van mening over de vraag of deze afspraken vooraf of tijdens het gesprek moeten worden gemaakt. Argumentatief gezien zijn beide manieren toegestaan, op

voorwaarde dat er duidelijkheid is over de procedure en beide partijen er-mee instemmen.

Ten slotte dient er tijdens de openingsfase ook overeenstemming te worden bereikt over de materiële uitgangspunten van de discussie. Over welke feiten, waarden en criteria zijn de partijen het eens? Zoals we in de literatuur hebben gezien, is het doorgaans de arts die de uitgangspunten bepaalt voor zover deze betrekking hebben op de medische feiten, en is het de patiënt die de uitgangspunten bepaalt die betrekking hebben op zijn eigen voorkeuren en waarden. Zowel de arts als de patiënt worden in principe geacht de door de andere partij ingebrachte uitgangspunten te accepteren. Deze taakverdeling bij het vaststellen van uitgangspunten vloeit voort uit de deontische en epistemische asymmetrieën die wij in de tweede paragraaf hebben uiteengezet.

Argumentatiefase

Het doel van de argumentatiefase is de aanvaardbaarheid van het standpunt dat ter discussie staat te toetsen door middel van een uitwisseling van argumenten en kritiek (Van Eemeren & Grootendorst, 2004, p. 61). In het proces van SDM gaat het erom een keuze te maken uit verschillende behandelopties. We zullen de elementen die bij dit beslissingsproces een argumentatieve rol spelen hieronder langslopen.

Ten eerste vindt het maken van een keuze altijd plaats op basis van een of meer beslissingscriteria. In het proces van SDM zijn deze criteria gerelateerd aan de consequenties van het uitvoeren van de verschillende behandelingen. Voorbeelden van dergelijke criteria zijn de mate waarin de behandeling tot levensverlenging leidt en de aard en intensiteit van de bijwerkingen. Omdat van de arts verwacht mag worden dat hij (meer dan de patiënt) over de vereiste biomedische kennis over deze consequenties beschikt, behoort het tot de argumentatieve taken van de arts om deze in het gesprek naar voren te brengen.

Zoals eerder beschreven valt te verwachten dat de patiënt de biomedische consequenties van de behandelopties niet allemaal even belangrijk vindt. Een tweede element dat in het proces van SDM een rol speelt is daarom een bepaalde beslissingsregel, die tot uitdrukking brengt hoe de scores van de behandelopties op de beslissingscriteria moeten worden gewogen. Idealiter reflecteert zo'n beslissingsregel welke waarden voor de patiënt het meest van belang zijn. Het kan bijvoorbeeld zijn dat de patiënt een substantiële levensverlenging belangrijker vindt dan de last die hij van bepaalde bijwerkingen ondervindt of juist andersom. Omdat de patiënt als

enige kan bepalen wat zijn eigen preferenties zijn en hoe deze zich tot elkaar verhouden, behoort het tot de argumentatieve taken van de patiënt om deze beslissingsregel in het gesprek naar voren te brengen, hierbij al dan niet geholpen door de arts.

In onze pragma-dialectische reconstructie van de argumentatiefase in het proces van SDM fungeren de medische consequenties en de beslissingsregel als argumenten ter ondersteuning van de uiteindelijke keuze voor de behandeloptie. Preciezer gezegd wordt het standpunt wat de beste behandeloptie is verdedigd door de volgende argumentatie:

standpunt	Behandeloptie B_i is de beste keuze uit de medisch acceptabele behandelopties B_1 - B_n
argument	Behandeloptie B_i is van de medisch acceptabele opties het meest in overeenstemming met voorkeur V_i van de patiënt
beslisregel	De behandeloptie die het meest in overeenstemming is met voorkeur V_i van de patiënt is de beste keuze uit de medisch acceptabele behandelopties B_1 - B_n

Afsluitingsfase

Het doel van de afsluitingsfase is vast te stellen wat de uitkomst van de discussie is (Van Eemeren & Grootendorst, 2004, pp. 61-62, 154). Is het verschil van mening opgelost en zo ja, ten gunste van wie? In plaats van opgelost, kan een geschil ook beslecht worden. In dat geval zijn de partijen het niet echt op basis van argumenten eens geworden, maar wordt er een andere manier gekozen om de knoop door te hakken.

Volgens de literatuur over SDM zou het verschil van mening tussen arts en patiënt idealiter opgelost en niet beslecht moeten worden. Zowel de arts als de patiënt zouden er aan het eind van het proces van overtuigd moeten zijn dat een bepaalde behandeling de beste keuze is. Omdat de patiënt in principe het beslisrecht heeft, kan deze ook wanneer de arts een bepaalde voorkeur heeft, toch voor een andere behandeling kiezen. Als er meerdere gelijkwaardige opties zijn, zou de arts het met die keuze op institutionele gronden eens moeten zijn (de keuze is dan immers subjectief en de patiënt heeft het recht om de beslissing te nemen). Ook als er volgens de arts een behandeling is die op medische gronden duidelijk de voorkeur heeft, heeft de patiënt toch het recht om voor een andere behandeling te kiezen.³ Dat zou betekenen dat het geschil beslecht wordt, of dat de arts zich terugtrekt en de patiënt naar een andere arts verwijst.

Conclusie

In dit artikel hebben we een pragma-dialectische vertaling gegeven van het proces van shared decision-making (SDM) die kan dienen als een analytisch instrument voor het reconstrueren van de discussiebijdragen van arts en patiënt aan het beslissingsproces in medische consultaties. Onze reconstructie van het proces van SDM is in principe op alle medische consultaties van toepassing; ook al kan de aard van de overwegingen die tot een uiteindelijke beslissing leiden per type consultatie variëren, er zal steeds sprake zijn van een combinatie van medische evidentie en voorkeuren van de patiënt. Wel worden in de literatuur over SDM contexten genoemd die bij uitstek vragen om SDM, omdat de opvattingen en de ervaringen van de patiënt daarbij van doorslaggevend belang worden geacht (zoals bijvoorbeeld in de palliatieve of chronische zorg).

Door het proces van SDM inclusief de verschillende invullingsvarianten te vertalen in pragma-dialectische termen wordt duidelijk wat in de verschillende fasen van het beslissingsproces de argumentatieve taken zijn van arts en patiënt en wat voor variatie hierin mogelijk is afhankelijk van factoren als de aard van de beslissing die ter discussie staat en de voorkeuren van de patiënt. Nog belangrijker is dat met een dergelijke analyse het noodzakelijke voorwerk is gedaan voor een preciezere evaluatie van het besluitvormingsproces. Tot nog toe beperkte de beoordeling van het medische besluitvormingsproces zich in de vakliteratuur tot de vraag of alle essentiële elementen van SDM wel aanwezig zijn en of artsen niet te sturend of persuasief optreden. Dit laatste werd in de praktijk van geval tot geval, op basis van niet duidelijk gearticuleerde normen vastgesteld. Door gebruik te maken van de in het pragma-dialectische model van een kritische discussie geformuleerde normen voor het voeren van een redelijke discussie, kan niet alleen op een meer gedetailleerde, maar ook op een meer verantwoorde manier worden nagegaan of het proces van SDM in concrete gevallen op redelijke wijze verloopt.

Noten

1. Dit recht vloeit voort uit het recht op *informed consent*, wat inhoudt dat een behandeling alleen mag worden uitgevoerd als de patiënt hier toestemming voor heeft gegeven na zorgvuldig geïnformeerd te zijn over de voor- en nadelen ervan. In Nederland is dit recht vastgelegd in de Wet op de Geneeskundige BehandelingsOvereenkomst (WGBO).
2. Het kan ook de patiënt zijn die het standpunt verdedigt dat een bepaalde behandeling medisch gezien de beste keuze is.

3. Lindström en Weatherall noemen dit recht van de patient '[the] fundamental patient's right to reject treatment proposals regardless of the scientific evidence that undergirds such proposals' (2015, p. 49).

Bibliografie

- Charles, C., Gafni, A., & Whelan, T. (1997). Shared decision-making in the medical encounter. What does it mean? (or it takes at least two to tango). *Social Science & Medicine*, 44(5), 681-692.
- Charles, C., Gafni, A., & Whelan, T. (1999). Decision-making in the physician-patient encounter. Revisiting the shared treatment decision making model. *Social Science & Medicine*, 49(5), 651-661.
- Dubov, A. (2015). Ethical persuasion. The rhetoric of communication in critical care. *Journal of Evaluation in Clinical Practice*, 21, 496-502.
- Eemeren, F. H. van, & Grootendorst, R. (1984). *Speech Acts in Argumentative Discussions*. Dordrecht/Cinnaminson: Foris Publications.
- Eemeren, F. H. van, & Grootendorst, R. (2004). *A Systematic Theory of Argumentation. The Pragmadiadical Approach*. Cambridge: Cambridge University Press.
- Emanuel, E. J., & Emanuel, L. L. (1992). Four models of the physician-patient relationship. *Journal of the American Medical Association*, 267(16), 2221-2226.
- Engelhardt, E., Pieterse, A. H., Van der Hout, A., De Haes, H. J. C. J. M., Kroep, J. R., Quarles van Ufford-Mannasse, P., Portielje, J. E. A., Smets, E. M. A., & Stiggelbout, A. M. (2016). Use of implicit persuasion in decision making about adjuvant cancer treatment. A potential barrier to shared decision making. *European Journal of Cancer*, 66, 55-66.
- Goodnight, T. G. (2006). When reasons matter most: Pragma-dialectics and the problem of informed consent. In P. Houtlosser & M. A. van Rees (Eds.), *Considering Pragma-dialectics. A Festschrift for Frans H. van Eemeren on the occasion of his 60th birthday* (pp. 75-85). Mahwah, N.J.: Lawrence Erlbaum Associates.
- Ishikawa, H., Hashimoto, H., & Kiuchi, T. (2013). The evolving concept of 'patient-centeredness' in patient-physician communication research. *Social Science & Medicine*, 96, 147-153.
- Kaba, R., & Sooriakumaran, P. (2007). The evolution of the doctor-patient relationship. *International Journal of Surgery*, 5, 57-65.
- Karnieli-Miller, O., & Eisikovits, Z. (2009). Physician as partner or salesman. Shared decision making in real-time encounters. *Social Science & Medicine*, 69, 1-8.
- Klaveren, H. van (2017). *Evaluating the applicability of the shared decision-making model. The asymmetry between doctor and patient in the 21st century. Masterscriptie*, Universiteit van Amsterdam.
- Labrie, N. H. M. (2014). *For the sake of argument. Considering the role, characteristics and effects of argumentation in general practice consultation. Dissertation*, University of Lugano.
- Laidsaar-Powell, R. C., Butow, P. N., Bu, S., Charles, C., Gafni, A., Lam, W. W. T., Jansen, J., McCaffery, K. J., Shepherd, H. L., Tattersall, M. H. N., & Juraskova, I. (2013). Physician-patient-companion communication and decision-making. A systematic review of triadic medical consultations. *Patient Education and Counseling*, 91, 3-13.
- Lindström, A., & Weatherall, A. (2015). Orientations to epistemics and deontics in treatment discussions. *Journal of Pragmatics*, 78, 39-53.
- Makoul, G., & Clayman, M. L. (2006). An integrative model of shared decision making in medical encounters. *Patient Education and Counseling*, 60, 301-312.

- Raymond, C. W., (2014). Conveying information in the interpreter-mediated medical visit. The case of epistemic brokering. *Patient Education and Counseling*, 97, 38-46.
- Snoeck Henkemans, A. F., & Mohammed, D. (2012). Institutional constraints on strategic maneuvering in shared decision medical decision making. *Journal of Argumentation in Context*, 1(1), 19-32.
- Stevanovic, M. (2013). *Deontic rights in interaction. A conversation analytic study on authority and cooperation. Dissertation*, University of Helsinki.
- Stiggelbout, A. M., Pieterse, A. H., & De Haes, J. C. J. M. (2015). Shared decision making. Concepts, evidence, and practice. *Patient Education and Counseling*, 98, 1172-1179.
- Towle, A., & Godolphin, W. (1999). Framework for teaching and learning informed shared decision making. *British Medical Journal*, 319(7212), 766-769.

Blijvende impact

Onderzoek robuust en relevant maken

Catherine E. Snow

Harvard Graduate School of Education

TVT 41 (1): 295–307

DOI: 10.5117/TVT2019.1.020.SNOW

Abstract

Lasting impacts: Making research robust and relevant

Lessons gleaned from educational research become sustainable only if incorporated into standard practice, institutionalized by being integrated into pedagogical approaches and teacher preparation, and recognized as the default way of doing things. The rarity with which research findings have had this sort of consequence in the field of education constitutes a major challenge to the field. One approach to bridging this research-practice gap is the mechanism of research-practice partnerships (RPPs), designed to ensure optimal relevance of research to problems of practice. This paper describes SERP (Strategic Education Research Partnership), one of the earliest-established RPPs in the U.S. and, more specifically, Word Generation, a program that addresses the literacy needs of middle school students. This example demonstrates the affordances of the RPP-model.

Keywords: sustainability, educational research, research-practice partnerships, literacy problems, vocabulary instruction, classroom discussion

Duurzaamheid in onderwijsonderzoek

Het begrip duurzaamheid met betrekking tot onderzoek heeft vele dimensies. Het omvat in ieder geval de dimensies levensduur en bestendigheid en dat impliceert, zeker in het licht van de huidige crisis in de sociale wetenschappen, de dimensie replaceerbaarheid. Vooral op het gebied van

onderwijsonderzoek hebben duurzame en repliceerbare resultaten echter niet altijd langdurige effecten in de praktijk. De duurzaamheid van onderwijsonderzoek zit in de aanpassingen die worden doorgevoerd in de aanpak van onderwijs. Met andere woorden, de lessen die we kunnen leren uit onderzoek worden alleen duurzaam als ze worden opgenomen in de standaardpraktijk, geïnstitutionaliseerd door integratie in pedagogische benaderingen en opleidingen van docenten, en erkend als de standaardmethode. Het komt zelden voor dat onderzoeksresultaten dergelijke gevolgen hebben in het onderwijs. Dit vormt een grote uitdaging voor het vakgebied.

Docenten worden daarom geconfronteerd met twee blijvende problemen: onderzoeksresultaten worden zelden vertaald naar verbeteringen in de praktijk en veranderingen in de praktijk hebben zelden een sterke wetenschappelijke basis. In feite zijn dit twee kanten van dezelfde medaille: onderzoek is op de een of andere manier moeilijk om te zetten in kennis die relevant is voor de praktijk en de praktijk schenkt te weinig aandacht aan richtlijnen uit onderzoek.

Vanuit de onderzoeker bezien kunnen we opmerken dat er geen sterke hefboomen zijn om de praktijk te beïnvloeden en dat scholen onvoldoende vermogen hebben om de (soms behoorlijk subtiële of ingewikkelde) lessen uit onderzoek over te nemen en te gebruiken. Als we het daarentegen vanuit de praktijk bezien, valt op dat er geen stimulans is om te leren van onderzoek, dat praktijkbeoefenaars niet voldoende toegerust zijn om wetenschappelijke kennis te onderzoeken, gebruiken en toe te passen en dat er geen robuuste industrie bestaat van uitgevers en ontwikkelaars van professioneel materiaal die tegemoet kunnen komen aan de vraag van docenten naar eenvoudige boodschappen en snelle oplossingen.

De discussie over dit tweeledige probleem woedt al decennia en er zijn meerdere mechanismen voorgesteld om de situatie te verbeteren. In de jaren '80 verschenen in de VS met steun van het Ministerie van Onderwijs handboeken over 'onderzoek naar praktijk', met het idee dat het vergemakkelijken van communicatie tussen onderzoekers en praktijkbeoefenaars het probleem zou oplossen. Anderen hebben voorgesteld dat de kloof tussen onderzoekers en praktijkbeoefenaars verkleind moet worden door het stimuleren van actieonderzoek, waarbij docenten en leidinggevers zelf als onderzoekers optreden in hun werkomgeving. Weer een andere aanpak die in de VS veel wordt toegepast is het ontwikkelen van prestatienormen om op onderzoek gebaseerde richtlijnen direct relevant te maken voor docenten. Helaas ging men er in de VS vanuit dat die normen gedrag alleen zouden veranderen in de context van een 'high-stakes accountability system', wat weer een lange lijst aan onfortuinlijke gevolgen veroorzaakte (Koretz, 2008,

2017). In andere onderwijssystemen worden normen ondersteund door een benadering gericht op het opbouwen van vermogen, waarbij scholen die bij de inspectie tekortschieten meer ondersteuning krijgen – een vriendelijker aanpak, maar ook een die vaak uitgevoerd wordt zonder het voordeel van nauwgezette begeleiding vanuit de wetenschap.

Onderzoek-praktijksamenwerkingen

Een heel andere aanpak voor het overbruggen van de kloof tussen onderzoek en praktijk zijn de ‘research-practice partnerships’ (RPPs), ontworpen om een optimale relevantie van onderzoek voor praktijkproblemen te waarborgen. Hoe werken zulke RPPs? Een van de basisprincipes is dat praktijkbeoefenaars en onderzoekers als gelijken samenwerken, waarbij iedere groep unieke expertise te berde brengt. Op basis van dit uitgangspunt zijn er vele modellen mogelijk voor RPPs. Ze kunnen klein zijn, kort duren en gericht zijn op een specifiek probleem, of groot, langdurig en veelomvattend in hun aanpak van problemen. Ze kunnen op een plek plaatsvinden, of verdeeld over locaties. Ze kunnen zich richten op het verbeteren van administratieve structuren of van de praktijk tijdens de lessen. Enkele gelukkige RPPs beschikken over kernfinanciering die hen in staat stelt om gericht te blijven op een centrale missie, terwijl andere meer financiële obstakels tegenkomen en afhankelijk zijn van projectfinanciering die het werk zijpaden op kan sturen.

Een aantal voorbeelden zal wellicht verduidelijken hoe uiteenlopend de constructies zijn die onder de noemer RPP vallen. Het Chicago Consortium for School Research (CSSR; <https://consortium.uchicago.edu/>), een eerbiedwaardig en duurzaam model, is voornamelijk data-georiënteerd: het richt zich op het verzamelen en gebruiken van data over patronen van voortgang bij leerlingen, om het district te kunnen informeren over de gevolgen van verschillende praktijken en om de gevolgen van aanpassingen aan de praktijk bij te houden. De CSSR kenmerkt zich door gecentraliseerd onderzoekspersoneel en regelmatige rapporten. Het andere uiterste is REX (Research Expertise Exchange), een informele online organisatie van Ierse geletterdheidsonderzoekers en -docenten die bedoeld is om directe lijnen van communicatie tussen hen te faciliteren. Of zoals een van de oprichters van REX het uitlegt, “I suppose you could say that it is a cross between Facebook and Tinder for teacher research!” (<https://youtu.be/qYfg1LAKqM>)

Zowel het Amerikaanse Ministerie van Onderwijs als de Spencer Foundation, de grootste filantropische steunpilaar van onderwijsonderzoek

in de VS, hebben de afgelopen jaren RPPs gefinancierd. Onder de door hen ondersteunde initiatieven vallen samenwerkingen tussen openbare scholen en non-profitorganisaties die zijn opgericht om hen te ondersteunen, zoals de onderwijsonderzoeksconsortia in Baltimore (www.baltimore-berc.org) en Philadelphia (www.phillyeducationresearch.org); samenwerkingen tussen netwerken van scholen of districten en onderzoeksorganisaties; samenwerkingen tussen universiteiten en scholencomplexen; en samenwerkingen waarbij de praktijkbeoefenaars in gemeenschap gevestigde organisaties of nationale dienstverleners waren. Het moge duidelijk zijn dat er onder deze omstandigheden een veelheid aan modellen voor het opzetten van RPPs is.

Een van de eerste RPPs in de VS is de SERP (Strategic Education Research Partnership), die bovendien een bijzonder goed uitgewerkt model en verzameling van principes heeft. SERP werd opgericht onder toezicht van de National Academy of Sciences in de vroege jaren '60, met steun van de toenmalige voorzitter Bruce Alberts. Alberts, zelf moleculair bioloog, ergerde zich eraan dat de onderwijspraktijk niet vooruitgebracht of verbeterd werd door de producten van onderwijsonderzoek, in tegenstelling tot de gezondheidszorg, landbouw en verkeersplanologie. Geïnspireerd door de aanpak in deze vakgebieden droeg hij zorg voor de oprichting van een commissie die ongebonden en creatief moest nadenken over onderwijsonderzoek. Een van de kenmerken van medisch, landbouwkundig en verkeerstechnisch onderzoek was de nauwe samenwerking tussen onderzoekers en beoefenaars: kankeronderzoekers waren vaak ook beoefenaars in laboratoria in een klinische setting; landbouwonderzoekers huurden lokale boeren in om samen nieuwe zaadsoorten, pesticiden en meststoffen op aangrenzende gebieden uit te testen en de resultaten bij te houden. De afstand tussen onderzoek en praktijk was veel kleiner dan in het onderwijs, waar onderzoekers vaak weinig connecties hadden met scholen en waar de onderzoekers die wel betrokken waren bij scholen (bijvoorbeeld in de toezichthouding op docententraining) meestal geen onderzoek deden. Er werd nu ingezet op het opbouwen van een nieuwe relatie tussen onderzoekers en beoefenaars die het probleemoplossende vermogen van onderzoek zou vergroten en ervoor zou zorgen dat onderzoek meer kennis vergaart over praktijksituaties die kan bijdragen tot verbeteringen in onderwijs (Donovan, Wigdor, & Snow, 2003).

Daarvoor ontwikkelde SERP een verzameling uitgangspunten die erop gericht zijn om het functioneren van de samenwerkingsverbanden tussen onderzoek en onderwijs te optimaliseren en de relevantie van onderzoek

voor onderwijsresultaten te maximaliseren. Het eerste uitgangspunt hield in dat de samenwerkingsverbanden zodanig ontworpen moesten zijn dat ze duurzaam en voor beide partijen nuttig waren. Een tweede uitgangspunt was dat het werk in ieder samenwerkingsverband moest beginnen met een urgent praktijkprobleem, aangedragen door een van de partners uit het onderwijs. Dat probleem kon verder toegespitst, uitgebreid of aangepast worden om ervoor te zorgen dat het grondig aangepakt kon worden, maar het basisdoel bleef het oplossen van het door de praktijkbeoefenaars aangestipte probleem.

SERP werkte dus via een systeem van netwerken van werklocaties waar een school of district steeds zelf hun urgente praktijkprobleem definieerde. SERP rekruteerde vervolgens een groep onderzoekers met relevante expertise om samen met vertegenwoordigers van de werklocatie tot een oplossing te komen. Tijdens de samenwerking komen vaak nieuwe problemen bovendrijven, die SERP dan weer door het onderzoeksteam kan laten bekijken. SERP heeft inmiddels bewezen dat het in staat is om met praktijkbeoefenaars samen te werken om belangrijke vragen te formuleren, om patronen in resultaten te begrijpen, om concrete tools te ontwikkelen op basis van de resultaten en om de verschillen in getrouwheid van toepassing en in effectiviteit als bron te gebruiken om de tools te verbeteren. Met andere woorden, SERP heeft zich vooral gericht op het bewerkstelligen van veranderingen in de standaardpraktijken in het klaslokaal. SERP werkt nauwe samen met administratoren en docenten in samenwerkingsdistricten om dat doel te bereiken, maar ontwikkelt ook oplossingen met het oog op een bredere relevantie.

Sinds zijn oprichting meer dan tien jaar geleden heeft SERP een brede verscheidenheid aan tools ontwikkeld, getest en afgeleverd die ontworpen zijn om docenten te ondersteunen in het gebruik van effectieve methodes in de klas. SERP heeft een twaalfstal tools ontwikkeld met verschillende focuspunten (zie <http://serpinstitute.org/> en Donovan, Snow, & Daro, 2013 voor meer informatie):

- Verbeteren van wiskundeonderwijs (Algebra by Example, Poster Problems, 5x8 Card)
- Toetsen van geletterdheidsvaardigheden van 'middle grade'-scholieren (RISE)
- Evalueren en promoten academische taalvaardigheden (CALS-I)
- Behouden van aandacht studenten (Focus Five)
- Stimuleren geavanceerde geletterdheids- en redeneervaardigheden (Word Generation, STARI, Reading to Learn in Science Website).

Een kernelement van de SERP-aanpak is de gelijktijdige aandacht voor de samenhangende behoeftes enerzijds van studenten aan toegang tot leer-mogelijkheden en anderzijds van docenten aan ondersteuning bij het bieden van die mogelijkheden, evenals de behoefte van scholen aan hulp bij het opbouwen van het organisationele vermogen dat nodig is om uitgebreid leren te implementeren.

SERP en de Boston Public Schools

SERP ging zijn eerste samenwerkingsverband aan in 2003 met de Boston Public Schools (BPS). De toenmalige BPS-voorzitter Thomas Payzant identificeerde lage niveaus van tekstbegrip bij leerlingen in 'middle school' ('grades' 6-8, scholieren van tussen de 10 en 14 jaar oud) als het urgente probleem waarvoor hulp nodig was. Normaal gesproken is er weinig expliciet geletterdheidsonderwijs in 'grades' 6-8, omdat ervan uit wordt gegaan dat de kinderen in 'grades' 1-3 hebben leren lezen en in 'grades' 4 en 5 hun taal-kennis en tekstbegrip verder hebben ontwikkeld. In werkelijkheid lopen scholieren in 'middle school' vaak tegen nieuwe obstakels aan, vooral bij het verwerken van en leren uit gamma- en bètawetenschappelijke teksten. Omdat het onderwijs wordt gesplitst in departementen vanaf de 'middle school', beschikken docenten van de gammavakken en de bètavakken vaak niet over de vaardigheden om deze tekstbegripsproblemen te begrijpen of te verhelpen. In plaats daarvan lossen ze het probleem vaak op door hun leerlingen te vertellen wat er in het boek staat.

Een eerste taak van het onderzoeksteam was dan ook om vast te stellen hoe de docenten tegen dit geletterdheidsprobleem aankeken. Waren ze het ermee eens dat de leerlingen problemen hadden met tekstbegrip en wat zagen zij als de voornaamste obstakels? Een jaar lang observeren in klaslokalen en interviewen van docenten resulteerde in een lange lijst van hypothesen over de redenen waarom scholieren moeite hadden met tekstbegrip. Docenten noemden onder andere lage motivatie, gebrek aan uithoudingsvermogen en slechte inferentievaardigheden, maar het meest genoemde probleem was de ontoereikendheid van hun woordenschat – niet de specifieke vocabulaire van de wetenschap of geschiedenis, maar algemene academische woorden waarvoor geen van de docenten zich geroepen voelde om die in zijn lessen te verwerken. Op basis van het algemene principe dat we oplossingen moesten voordragen voor problemen waar men zich zorgen over maakte, stelden we vast dat onze eerste taak in BPS

het ontwikkelen van een programma voor academische vocabulaire voor de 'middle school' was.

De specifieke kenmerken van Word Generation, het programma dat we ontwikkelden, kwamen uit twee bronnen: de kennis vergaard in decennia-lang onderzoek naar de kenmerken van succesvol vocabulaireonderwijs en de praktische beperkingen van het werken in scholen. De wetenschap stelde de volgende kenmerken voor:

- leerling frequent in aanraking brengen met beoogde woorden;
- mogelijkheden bieden om de woorden te horen en lezen in een grote verscheidenheid aan rijke semantische contexten;
- mogelijkheden bieden om de woorden te gebruiken in authentieke gesproken en geschreven communicatie;
- bieden van scholievriendelijke definities nadat leerling het woord een aantal keer is tegengekomen;
- bieden van metalinguïstieke strategieën, zoals het uitleggen van morfologische structuren en het identificeren van cognaten (vooral voor leerlingen die een tweede taal leren).

De meeste kenmerken waren op relatief gemakkelijke wijze op te nemen in het curriculaire materiaal, met uitzondering van de eis dat de leerlingen de nieuwe woorden zelf gebruiken. Hoe zorg je ervoor dat leerlingen specifieke woorden op een authentieke manier gebruiken? De standaardpraktijk bij het leren van nieuwe woorden is dat de leerling gevraagd wordt om een zin te formuleren met de beoogde woorden, wat niet bepaald authentiek is. We kwamen op het idee om het curriculum in de vorm te gieten van een reeks discussieonderwerpen die het mogelijk maken om normale debatten te voeren waarin de woorden gebruikt konden worden. Daarvoor overlegden we met docenten om een lijst met onderwerpen samen te stellen die hun leerlingen interessant zouden vinden en ontwikkelden we rond ieder van deze onderwerpen lesmateriaal voor een week, bijv. *Zou een gestandaardiseerde toets een vereiste moeten zijn voor het behalen van je middelbare schooldiploma? Zouden scholen positieve discriminatie moeten toepassen? Zouden scholen of ouders verantwoordelijk moeten zijn voor seksuele voorlichting? Is het juist dat professionele sporters toplonen verdienen? Zouden leerlingen betaald moeten worden voor prestaties op school?* Om ervoor te zorgen dat de nieuwe woorden in verschillende semantische contexten voorkwamen en om te benadrukken dat academische woordenschat een gedeelde institutionele verantwoordelijkheid was, ontwierpen we ook een reeks activiteiten die verdeeld waren over alle centrale vakgebieden: we introduceerden het onderwerp in de lessen Engels op maandag en breidden

het onderwerp uit met activiteiten voor natuurkunde en wiskunde, debatten voor de gammavakken en het schrijven van een argumentatief essay op vrijdag (zie Snow & Lawrence, 2011, voor een overzicht van dit werk).

Ondertussen kregen we ook aanzienlijke begeleiding van praktijkbeoefenaars van BPS over specifieke kenmerken van het programma. De docenten die vroege versies uittesten gaven aan dat zij hun leerlingen vijf woorden per week konden leren, niet de door ons voorgestelde tien. Ze onderzochten ook mogelijke onderwerpen en gaven feedback over het soort wiskundeoefeningen dat zij bruikbaar vonden. We werden geconfronteerd met de realiteit van lesstofplanning, waardoor er bij wiskunde, natuurkunde en de gammavakken slechts weinig tijd over was voor activiteiten die niet als kernmateriaal gezien werden. Uiteindelijk ontwikkelden we een programma van 20 minuten per dag dat roteerde tussen de kernvakken, zodat slechts 20 minuten per week werd gevraagd van de lessen in wiskunde, natuurkunde en de gammavakken, en slechts 40 minuten van de lessen Engels.

De belangrijkste instructieve activiteiten voor elke dag in de reeks, behalve de schrijfopdracht op de vrijdag, waren discussies in kleine of grotere groepen. Het introductie-essay, de wiskundeproblemen, de natuurkundevragen en de discussies bij de gammavakken formuleerden we allemaal als mogelijkheden voor de leerlingen om samen te werken en voor de docenten om debatten te leiden in plaats van antwoorden te geven. Voor iedere curriculaire eenheid selecteerden we dilemma's die zich goed voor discussie leenden. Daarbij hadden we nog niet door wat zulke vragen zonder correcte antwoorden aan mogelijkheden boden met betrekking tot het creëren van nieuwe structuren in het klaslokaal: plotseling is de docent niet meer de autoriteit en krijgen leerlingen de gelegenheid om hun eigen standpunten te ontdekken en verdedigen. Hiermee ontstaat een situatie die nogal afwijkt van de norm. Zelfs waar discussie al onderdeel was van het lesprogramma, werden de leerlingen vaak nog naar het 'beste' antwoord gestuurd, in plaats dat hun toegestaan werd om hun eigen standpunten te ontwikkelen en op elkaars argumenten te reageren.

Word Generation werd goed ontvangen door onze partners van de Boston schoolgemeenschap (Snow, Lawrence, & White, 2009). We werden gevraagd om genoeg curriculaire eenheden te ontwikkelen zodat het programma in een hele school gebruikt kon worden op basis van een driejarige cyclus. De 72 eenheden die daaruit voortkwamen, nu herzien en opnieuw ontworpen, zijn vrij beschikbaar voor docenten van over de hele wereld (wordgen.serpmedia.org). Ze zijn tot nu toe 30.000 keer gedownload door docenten uit alle vijftig deelstaten van de VS en meer dan tien

andere landen. Een gerandomiseerd clusteronderzoek van het programma toonde aan dat het succesvol was in het aanleren van de beoogde woorden (Lawrence, Rolland, Branum-Martin, & Snow, 2014), dat de behaalde resultaten zichtbaar bleven tot minstens een jaar na het programma (Lawrence, Capotosto, Branum-Martin, White, & Snow, 2012) en dat de resultaten van leerlingen van wie Engels niet de moedertaal was nog sterker waren (Hwang, Lawrence, Mo, & Snow, 2013; Lawrence, Crosson, Paré-Blagoev, & Snow, 2015). Toch lukte het ons niet om ons uiteindelijke doel te bereiken: het verbeteren van de leesvaardigheid van scholieren.

Het grootste effect van Word Generation betrof wellicht niet zozeer het gebied van de woordenschat van scholieren maar het gebruik van klassikale discussies. We ontwikkelden een observatieschaal voor klassikale discussies, voornamelijk om de getrouwheid van de implementatie te controleren (Lawrence et al., 2015). De kwaliteit van discussies in de Word Generation-klassen was .3-1.1 standaarddeviatie beter dan in controleklassen, met de grootste effecten bij wiskunde en natuurkunde. Tijdens de observaties werden we ons bewust van de kracht van de klassikale discussies die we observeerden: hun vermogen om leerlingen te activeren en kritisch denken te stimuleren. Docenten lieten ons weten dat ze zich nooit hadden gerealiseerd hoe bedachtzaam hun leerlingen waren en wat een goede ideeën ze hadden totdat ze begonnen met de discussies over morele en maatschappelijke vraagstukken. Om deze redenen begonnen we Word Generation meer te beschouwen als een discussieprogramma dan als een vocabulaireprogramma.

Word Generation, tweede generatie

De inzichten van de gebruikers van het originele Word Generation-materiaal en onze eigen observaties tijdens de lessen vormden een goede uitgangspunt voor ons om mee te dingen naar een ruime financieringsmogelijkheid vanuit het Amerikaanse Departement voor Onderwijs, genaamd Reading for Understanding. De financiering van Reading for Understanding stelde ons in staat om een evaluatie uit te voeren van een zeer uitgebreide en verrijkte versie van het programma dat begon in de vierde 'grade' in plaats van de zesde met 18 volledige weken aan lesmateriaal voor gamma- en bètavakken voor de zesde tot en met de achtste 'grade'. Bij het ontwerpen van de uitgebreide materialen vergrootten we de mogelijkheden voor en de begeleiding bij het gebruik van discussies tijdens de les. Ook testten we expliciet onze nieuwe theorie over de werking van het programma, namelijk dat het tekstbegrip

van 'middle grade'-scholieren afhankelijk was van drie vaardigheden die gestimuleerd werden bij discussies: academisch taalgebruik, het aannemen van verschillende perspectieven en argumentatie (LaRusso et al., 2016). Deze vaardigheden kwamen nu allemaal specifiek aan de orde in het curriculum.

Het uitgebreide Word Generation-programma bleek succesvol in het verbeteren van onze originele maatstaf, de beoogde vocabulaire, in het eerste jaar van de implementatie. In het tweede jaar waren de effecten meer verspreid (Jones et al., in press). We zagen significante effecten op de onderwezen vocabulaire, het perspectiefnemen, academisch taalgebruik (vooral voor leerlingen van wie Engels niet de moedertaal was) en tekstbegrip.

De sterkste effecten bleven echter beperkt tot de leerlingen in de vierde en vijfde 'grade' en vanzelfsprekend tot de leerlingen in de klassen met de meest grondige implementatie. Gemiddeld werd slechts rond de 50% van het curriculum daadwerkelijk gebruikt. De effecten in klassen waar minstens 70% van het curriculum werd gebruikt, waren sterk.

Het RPP-effect

Rest de vraag hoe dit tweedelige onderzoeksprogramma dat ik zojuist heb beschreven aantoont dat SERP of het onderzoek-praktijksamenwerkingsmodel (RPP) een belangrijke rol vervult. Ik zal me hier beperken tot de belangrijkste voordelen van het model. Ten eerste was het succes van het programma in zijn uiteindelijke vorm voor een groot deel te danken aan de samenwerking met docenten in onze eerste partnerscholen. Zij boden onmisbare begeleiding op het gebied van onderwerpen, woorden, verwachtingen en specifieke activiteiten. Ten tweede waren de evaluatieonderzoeken naar zowel de eerste als de tweede generatie van het programma grootschalige ondernemingen die samenwerkingen met meerdere school-districten vereisten om genoeg deelnemende scholen te krijgen. SERP heeft een veel betere uitgangspositie om de vele transacties en omvangrijke communicatie te organiseren die bij een dergelijk project komen kijken. Ten derde is het ontwikkelen van inhoudelijk materiaal voor een nieuw curriculum slechts één onderdeel van de totale taak van het ontwerpen van een curriculum dat bruikbaar is voor docenten tijdens echte lessen. Er zijn specifieke vaardigheden nodig voor het maken van een ontwerp dat scholieren aanspreekt en dat duidelijk aangeeft hoe de verschillende onderdelen samenhangen. SERP beschikt over een designteam dat ervaren is in zulke opdrachten en bereid ervoor te zorgen dat het curriculaire materiaal typografisch aantrekkelijk, feitelijk juist en up-to-date is. Ten vierde bestaat

het designteam onder anderen uit videomakers, die het gebruik van Word Generation tijdens lessen kunnen documenteren met het oog op outreach, professionele ontwikkeling en onderzoek. Ten vijfde, het designteam was ook verantwoordelijk voor het ontwikkelen en bijhouden van de website waarmee het Word Generation-materiaal werd verspreid en die professionele steun biedt voor toekomstige gebruikers. Ten zesde was SERP flexibel genoeg om door te blijven werken nadat de specifieke financiering was afgelopen, ondanks het feit dat het voornamelijk op projectbasis bekostigd wordt. Veelbelovende onderzoeksprojecten worden vaak afgekappt precies op het moment dat ze nuttige resultaten voor de praktijk opleveren, omdat universiteitsgebonden onderzoeksteams geen middelen hebben om het project voort te zetten aan het einde van de financieringsperiode. SERP biedt continuïteit voorbij het originele onderzoeksteam. Ten zevende, omdat SERP samenwerkt met een aantal districten en een verzameling goed bekendstaande tools heeft ontwikkeld, beschikt SERP over effectieve mechanismen voor het verspreiden van informatie over Word Generation en staat het bekend als een betrouwbare bron voor bruikbare en wetenschappelijk onderbouwde tools. Ten slotte is SERP als stabiele organisatie in staat om Word Generation uit te breiden door professionele ontwikkelingsinstituten aan te nemen en vergunningen uit te geven aan coaches. Het is lastig voor individuele onderzoekers om deze activiteiten zelf te ondernemen.

Kortom, het 15 jaar lange traject van Word Generation toont hoe de uitgangspunten ten grondslag aan een onderzoek-praktijksamenwerking zoals SERP functioneren. Het begon als antwoord op het urgente praktijkprobleem aangedragen door de Boston Public Schools, waar SERP een blijvende samenwerking mee heeft opgebouwd. Zoals verwacht bleek de oplossing voor het lokale probleem in Boston ook aantrekkelijk voor andere stedelijke gebieden, waar vaak hetzelfde probleem met geletterdheid onder jongeren speelde. Het ontwerp combineerde top-down kennis van jarenlang onderzoek over vocabulaireonderwijs met bottom-up kennis van docenten in de deelnemende scholen. Gedurende het project hebben we een scala aan tools ontwikkeld – curriculaire inhoud, docentenhandleidingen, online professionele ontwikkeling, doelopdrachten – en hebben we bovendien de kennis over leren en onderwijzen vergroot.

Conclusie

Toen Bruce Alberts de commissie voor onderwijsonderzoek oprichtte in 2000, legde hij hen twee vraagstukken voor:

- 1 Waarom heeft onderzoek wel bijgedragen aan innovatie en voortdurende verbeteringen op het gebied van gezondheidszorg, landbouw en transport, maar niet op het gebied van onderwijs?
- 2 Wat kunnen we daaraan doen?

Het commissierapport (Donovon, Wigdor, & Snow, 2003) beantwoordde de eerste vraag door erop te wijzen dat er veel nauwer en meer ononderbroken contact tussen onderzoekers en praktijkbeoefenaars is op het gebied van gezondheidszorg, landbouw en transport en dat er meer wederzijds respect is tussen de onderzoekers en praktijkbeoefenaars in deze vakgebieden dan in het onderwijs. Daarnaast zijn dit natuurlijk vakgebieden waarin veel geld wordt geïnvesteerd en wordt de expertise van zowel onderzoekers als beoefenaars op deze gebieden algemeen erkend door beleidsmakers en het publiek. Met betrekking tot het onderwijs daarentegen, wordt de kwaliteit van het onderzoek over het algemeen niet hoog ingeschat, denkt men dat docenten gemiddeld slecht voorbereid zijn op hun rol en vindt men iedereen die naar school is geweest een expert.

In reactie op de tweede vraag stelde het rapport van de SERP-commissie een goed gefinancierde onderneming voor de lange termijn voor, die de lessen en ontwikkelingen die plaatsvinden op diverse knooppunten in grote netwerken van lokale RPPs moet combineren. In werkelijkheid heeft SERP zes lokale knooppunten van onderzoek-onderwijs samenwerkingen opgericht die allemaal aan hun eigen urgente praktijkproblemen werken maar nauwelijks genoeg financiering of toekomstzekerheid voor de langere termijn hebben. Een onverwachte bijkomstigheid was echter dat het concept van RPPs wel is aangeslagen en dat er nu tientallen opereren in Amerikaanse scholen en onderwijsinstellingen. Hoewel die niet gelieerd zijn aan een centrale organisatie, kunnen ze wel een waardevolle bron vormen om te ontdekken wat de randvoorwaarden zijn voor succesvolle RPPs en welke omstandigheden hun duurzaamheid in gevaar brengen. Wellicht nog belangrijker dan de duurzaamheid van de RPPs is de duurzaamheid van de veranderingen die zij introduceren in de praktijk. Een centrale claim van het RPP-model is dat de innovaties in onderwijspraktijken die het gevolg zijn van de samenwerkingen tussen onderzoekers en praktijkbeoefenaars beter afgestemd zijn op gebruik tijdens de les en daarom duurzamer dan innovaties die worden voorgesteld door onderzoekers die niet hebben kunnen profiteren van een dergelijke samenwerking. Het lot van Word Generation en andere door SERP ontwikkelde tools op de lange termijn zal moeten uitwijzen of die claim terecht is.

Bibliografie

- Donovan, M. S., Snow, C.E., & Daro, P. (2013). *The SERP approach to problem-solving research, development, and implementation*. *National Society for the Study of Education Yearbook*, 112(2), 400-425.
- Donovan, M.S., Wigdor, S., & Snow, C.E. (Eds). (2003). *Strategic Education Research Partnership*. National Research Council, Washington DC: National Academies Press.
- Hwang, J. K., Lawrence, J., Mo, E., & Snow, C. (2015). Differential effects of a systematic vocabulary intervention on adolescent language minority students with varying levels of English proficiency. *International Journal of Bilingualism*, 19(3), 314-332.
- Jones, S. M., LaRusso, M., Kim, J., Kim, H Y, Selman, R., Uccelli, P., Barnes, S., Donovan, S., & Snow, C. (in press). Experimental effects of Word Generation on vocabulary, academic language, perspective taking, and reading comprehension in high poverty schools. *Journal of Research in Educational Effectiveness*.
- Koretz, D. (2008). *Measuring Up: What Educational Testing Really Tells Us*. Cambridge, MA: Harvard University Press.
- Koretz, D. (2017). *The Testing Charade: Pretending to Make Schools Better*. Chicago, IL: University of Chicago Press.
- LaRusso, M., Kim, H.Y., Selman, R., Uccelli, P., Dawson, T., Jones, S., Donovan, S., & Snow, C. (2016). Contributions of academic language, perspective taking, and complex reasoning to deep reading comprehension. *Journal of Research in Educational Effectiveness*, 9, 201-222.
- Lawrence, J., Capotosto, L., Branum-Martin, L., White, C., & Snow, C. (2012). Language proficiency, home-language status, and English vocabulary development: A longitudinal follow-up of the Word Generation program. *Bilingualism: Language and Cognition*, 15, 437-451.
- Lawrence, J., Crosson, A., Paré-Blagoev, J., & Snow, C. (2015). Word Generation randomized trial: Discussion mediates the impact of program treatment on academic word learning. *American Educational Research Journal*, 2(1), 1-37.
- Lawrence, J., Rolland, R., Branum-Martin, L., & Snow, C. (2014). Generating vocabulary knowledge for at-risk middle school readers: Contrasting program effects and underlying growth trajectories. *Journal of Education for Students Placed at Risk (JESPAR)*, 19, 76-97.
- Snow, C. E., & Lawrence, J. F. (2011). *Word Generation in Boston Public Schools: Natural history of a literacy intervention. Volume III, Senior Urban Education Research Fellowship Series*. Washington, DC: Council of Great City Schools.
- Snow, C., Lawrence, J., & White, C. (2009). Generating knowledge of academic language among urban middle school students. *Journal of Research on Educational Effectiveness*, 2, 325-344.

Ontnuchterende voorlichtingscampagnes over comazuipen bij studenten

Een experimenteel onderzoek naar de overtuigingskracht en effectiviteit van humor en angst, en inspelen op fysieke en sociale gevaren

Sara Verbrugge en Romane Morreale
KU Leuven

TVT 41 (1): 309–322

DOI: 10.5117/TVT2019.1.021.VERB

Abstract

Fear and humor appeals (with physical and social threats) on bingedrinking in student populations

What is the best way to make young adults aware of the costs of binge drinking? What is the most effective social marketing strategy to change a target audience's behavioral intention in a desired way? Fear appeals create an unpleasant tension, which will motivate the individual to reduce the uncomfortable feeling by adapting his behavior as recommended in the social marketing message. However, research exploring the effectiveness and the ethical framework of such fear appeals has yielded inconsistent findings. Therefore, humor is often suggested as an alternative strategy to influence behavioral intentions in health communication.

A 2 (appeal: fear vs. humor) x 2 (threat: physical vs. social) between-subjects online experiment was set up, with 164 participants (19-26 years old). Fear appeals appeared to persuade students to a greater extent compared to humor appeals. Similarly, physical threats were more effective than social threats. However, when combining those two message components frightful advertisements with a social threat led to the greatest message conviction. Furthermore, we recommend social marketers who are planning to use humor appeals in their anti-binge drinking campaign to focus on physical threats instead of social threats. The latter will only reduce the effectiveness of the message.

Keywords: health communication, fear appeals, physical and social threats, humor, persuasiveness, bingedrinking

Inleiding

Velen onder ons denken ongetwijfeld weemoedig terug aan wilde feestjes met de nodige alcoholconsumptie tijdens de studentenjaren. Problematisch wordt het pas als het drinken op dergelijke momenten telkens opnieuw de overhand neemt. We schetsen de problematiek van het comazuipen en geven aan waarom voorlichtingscampagnes nodig zijn. Vervolgens gaan we dieper in op sociale marketing, *fear appeals* die inspelen op fysieke of sociale gevaren, en de voordelen van humor in dergelijke campagnes. Daarna stellen we ons onderzoek voor, bespreken de resultaten en beschrijven de relevantie van de bevindingen in de discussie.

Probleemschets: comazuipen of *bingedrinken* bij studenten

Vaak wordt bingedrinken of comazuipen gedefinieerd als een hoge consumptie van alcohol in een korte periode. Zo kan er van bingedrinken worden gesproken bij het drinken van vijf of meer drankjes voor mannen en vanaf vier drankjes voor vrouwen op één enkele gelegenheid (Wolburg, 2001). Verschillende studies hebben aangetoond dat bingedrinken gevaarlijker is dan intermitterend alcoholgebruik (Stephens & Duka, 2008). Mogelijke gevolgen van bingedrinken zijn schade aan de organen, meer kans op een black-out, stijging van de bloeddruk, kans op een hartinfarct, leverkanker, borstkanker, alcoholvergiftiging, overgewicht, en verwondingen ten gevolge van dronken gewelddadigheid (Boles & Miotto, 2003; Hunter & Francescutti, 2013; Naimi, Nelson, & Brewer, 2010). Volgens de meeste empirische studies blijkt dit problematisch alcoholgebruik zich vooral voor te doen bij 18- tot 25-jarigen. Zo blijkt in de Verenigde Staten 75 à 90 procent van de studenten alcohol te drinken en hiervan blijkt er ongeveer 44 procent twee keer per week te bingedrinken (Wechsler, Davenport, Dowdall, Hoeykens, & Castillo, 1994). Bijgevolg is het relevant om geschikte voorlichtingsboodschappen te ontwikkelen waarin men studenten bewustmaakt van de mogelijke gevaren en risico's verbonden aan bingedrinken.

Literatuur: sociale marketing, *fear appeals* en humor

Bij socialemarketingcampagnes probeert men te bepalen welke strategie het beste zou werken om de gedragsintentie van een doelgroep op een positieve manier te beïnvloeden. Een klassiek onderwerp is het gebruik van *fear appeals* in voorlichtingscampagnes. Er bestaat onenigheid over de werking,

de effectiviteit en de ethische verantwoording van fear appeals (Hastings, Stead, & Webb, 2004; Leventhal, 1970; Witte, 1992). Als alternatief wordt humor vaak naar voren geschoven, maar deze strategie is voorlopig minder onderzocht (Yoon & Thinkham, 2013; Yoon, 2015).

Fear appeals: belang van angst en dreiging

In de context van socialemarketingboodschappen verwijst dreiging naar de inhoud van de boodschap en angst naar de potentiële emotionele respons op de boodschap (Cauberghe, De Pelsmacker, Janssens, & Dens, 2009). Angst vormt een basisemotie en kan worden gedefinieerd als een negatief gevoel, in combinatie met een hoge opwinding, dat tot stand komt door een significante en persoonlijk relevante dreiging (Easterling & Leventhal, 1989). De respons op een dreiging verschilt voor elke bevolkingsgroep en zelfs voor elk individu. Er bestaat geen optimaal niveau van fear appeal dat op alle individuen kan worden toegepast om het gewenste gedrag te verkrijgen (Latour & Zahra, 1989). Fear appeals worden dus het best telkens aangepast aan het specifieke segment van de bevolking voor wie de campagne bedoeld is (Burnett & Oliver, 1979).

Angst wordt als negatieve appeal frequenter gebruikt binnen sociale marketing in de hoop dat het opwekken van angst het gewenste gezondheidsgedrag tot gevolg zal hebben. Men vertrekt vanuit de idee dat fear appeals een emotioneel onevenwicht en spanning met zich meebrengen. Dat zorgt voor een onaangenaam gevoel, waardoor mensen gemotiveerd zullen zijn om het gewenste gedrag tot stand te brengen, dit oncomfortabel gevoel te reduceren en de dreiging weg te werken (Brennan & Binney, 2010; Easterling & Leventhal, 1989).

In de literatuur blijkt er veel inconsistentie te bestaan over de werking van fear appeals. Zo tonen sommige studies de effectiviteit van fear appeals aan (Stainback & Rogers, 1983), terwijl anderen net het omgekeerde aantonen (Kohn, Goodstadt, Cook, Sheppard, & Chan, 1982). Het is van belang dat fear appeals de juiste emoties oproepen en tegelijkertijd dat het individu gelooft dat hij/zij in staat is om deze dreiging op te lossen (Witte, 1992).

Fysieke en sociale fear appeals en hun effect op verschillende doelgroepen

We gaan dieper in op het onderscheid tussen fear appeals die een sociale dreiging ofwel een fysieke dreiging inhouden (Laroche, Toffoli, Zhang, & Pons, 2001; Schoenbachler & Whittler, 1996). Sociale en fysieke fear appeals verschillen in hoe het emotionele overtuigingsproces verloopt. Fysieke fear appeals houden rechtstreeks verband met onze menselijke sterfelijkheid,

terwijl sociale fear appeals verwijzen naar de mens en zijn zelfwaarde in een maatschappelijk perspectief (Martin & Kamins, 2010).

Fysieke fear appeals hebben betrekking op iemands lichaam, gezondheid en leven. Zo hanteert men in de context van drugsgebruik en het effect op het menselijk lichaam vaak beelden van jonge mensen die in het ziekenhuis liggen (Schoenbachler & Whittler, 1996). Ook in de anti-rookgezondheidscommunicatie benadrukt men vaak de fysieke gevolgen, zoals longkanker en snellere veroudering (Smith & Stutts, 2003). Fysieke fear appeals in een bingedrinkingcontext benadrukken vroegtijdige dood, risico op leverkanker en hartproblemen (Murgraff, Parrott, & Bennett, 1999).

Sociale fear appeals benadrukken de gevolgen van het gestelde gedrag voor onze sociale omgeving zoals sociale uitsluiting of schade berokkenen aan de gezondheid of het welzijn van een ander (bv. sigarettenrook inademen of iemand verwonden, Rayner, Baxter, & Ilicic, 2015). Het gebruik van sociale dreiging werd in het verleden minder onderzocht dan het gebruik van fysieke dreiging. Men ging er immers van uit dat verwondingen en confrontatie met de dood overtuigender waren dan sociale uitsluiting. Hier kwam de laatste jaren verandering in, omdat het gevaar voor sociale uitsluitingen voor bepaalde doelgroepen heel overtuigend is (Smith & Stutts, 2003). Voornamelijk tijdens de adolescentie wil men vriendschapsbanden sluiten met leeftijdsgenoten, wil men een onafhankelijke identiteit construeren, en lijkt men zich te willen onderscheiden van de ouders (Bloom, 1998). Volgens verschillende studies zouden sociale dreigingen dus het beste werken voor adolescenten (Schoenbachler & Whittler, 1996). Bovendien duiken fysieke dreigingen meestal pas na verloop van tijd op, waardoor veel adolescenten deze fysieke dreigingen negeren (Smith & Stutts, 2003). Echter, niet alle studies stemmen met deze visie overeen. Zo vonden Reardon en Miller (2008) net het omgekeerde effect, namelijk dat fysieke fear appeals effectiever zijn dan sociale fear appeals bij jongvolwassenen.

Humor in sociale marketing

Humor wordt vaak naar voren geschoven als alternatief om gedragsverandering teweeg te brengen, maar onderzoek is schaars (Lagerwerf, 2007; Yoon & Thinkham, 2013; Yoon, 2015a), zeker in de gezondheidscommunicatie (Hendriks & Janssen, 2018; Lewis, Watson, White, & Richard, 2007; Yoon, 2015b). Humor is nochtans een van de meest gehanteerde communicatiestrategieën in de reclamewereld (Lee & Lim, 2008). De effectiviteit van humor op aandacht, overtuiging, en houding ten opzichte van de boodschap

werd in het verleden vooral bestudeerd in commerciële reclamecampagnes (Speck, 1987; Weinberger & Gulas, 1992).

Ten eerste zou humor de aandacht van de ontvanger trekken en ook onze aandacht blijven vasthouden (Weinberger & Gulas, 1992). Anderzijds tonen andere studies de keerzijde, namelijk dat humor de aandacht van de inhoud van de boodschap wegneemt (Krishnan & Chakravarti, 2003).

Ten tweede zou humor ook een effect hebben op de overtuiging. Ook hier heerst er inconsistentie wat betreft de resultaten. Humor wekt positieve gevoelens op zoals blijdschap, speelsheid en vrolijkheid. Volgens sommigen zou humor leiden tot meer overtuiging van de boodschap omdat humor onze aandacht afleidt van de bedoeling om ons te overtuigen. Het oproepen van positieve emoties zou mogelijke ongewenste reacties, zoals de vorming van tegenargumenten, vermijden (Lee & Ferguson, 2002). Andere studies vonden dat humor geen toegevoegd effect had op overtuiging (Belch & Belch, 1984).

Ten derde zou humor onze houding ten opzichte van de campagne kunnen beïnvloeden. Indien men een positieve attitude of appreciatie ten opzichte van de communicatiebron wil ontwikkelen, raadt men aan om humor appeals te gebruiken. De positieve gevoelens over de advertentie kunnen vervolgens overgaan op het merk (Eisend, 2009). Zo zou het gebruik van humor de kans op defensieve reacties verminderen (Weinberger & Gulas, 1992). Daarentegen blijkt het gebruik van humor niet aanbevolen in serieuze en dringende situaties, omdat het de gepercipieerde ernst van de boodschap kan doen afnemen (Yoon, 2015a).

Onderzoek

In deze studie (zie ook Morreale, 2016) vergelijken we de effectiviteit van fear en humor appeals om studenten bewust te maken van de risico's verbonden aan bingedrinken. Verder gaan we na welke soort dreiging, namelijk een fysieke of sociale dreiging, het beste werkt, ook in combinatie met angstaanjagende of humoristische advertenties. Deze studie werd met andere woorden opgezet om de lacune aan empirisch onderzoek naar de effectiviteit van humor appeals in bedreigende boodschappen en de onenigheid over de effectiviteit van fear appeals te dichten. De sociaalmarketingbranche erkent immers wel het potentieel van beide strategieën, maar een vergelijking van beide strategieën bij het bingedrinkingsprobleem ontbreekt.

Onderzoeksvragen

Op basis van de literatuur stellen we volgende onderzoeksvragen op over attitude (1), over overtuigingskracht (2 en 3), en over de verwachte effectiviteit (4 en 5):

- 1 Leiden humoristische of angstaanjagende boodschappen tot een positievere attitude ten opzichte van de voorlichtingsadvertentie?
- 2 Welk dreiging (fysiek of sociaal) heeft de grootste overtuigingskracht?
- 3 Welke appeal (humor of angst) heeft de grootste overtuigingskracht?
- 4 Leiden humoristische of angstaanjagende boodschappen tot een grotere effectiviteit?
- 5 Welke soort boodschap (humor of angst) in combinatie met welke dreiging (fysiek of sociaal) leidt tot de grootste effectiviteit?

Methode

Deelnemers: Omdat verschillende studies erop gewezen hebben dat binge-drinken zich grotendeels voordoet tijdens de studentenjaren, hebben we onze doelgroep toegespitst op 18- tot 26-jarigen. 164 Vlaamse jongvolwassenen namen deel (gemiddelde leeftijd = 22 jaar, $SD = 1.3$). De steekproef bestaat voor 68.9 procent uit vrouwen ($n = 113$) en voor 31.1 procent uit mannen ($n = 51$).

Design: Het online experiment is een 2 (appeal: angst vs. humor) x 2 (dreiging: fysiek vs. sociaal) between-subjects design. Zoals hierboven beschreven, werden fysieke fear appeals geselecteerd die betrekking hebben op iemands lichaam, gezondheid en leven. Sociale fear appeals gaan over de gevolgen van bepaald gedrag voor onze sociale omgeving zoals sociale uitsluiting of schade berokkenen aan de gezondheid of het welzijn van een ander. Elke respondent werd aan de hand van het online platform Qualtrics op toevallige wijze toebedeeld aan een bepaalde conditie en werd gevraagd drie advertenties aandachtig te bekijken en te evalueren aan de hand van een gestandaardiseerde vragenlijst.

Materiaal: Er werden vier condities onderscheiden. De eerste conditie bevat boodschappen die angst oproepen door middel van een fysieke dreiging (de dood, zelf overreden worden door zich dronken in het verkeer te begeven en het slachtoffer zijn van dronken gewelddadigheid). De bedoeling van die boodschappen is dus om angst op te roepen bij de respondent en de fysieke gevolgen van binge-drinken te benadrukken. De tweede conditie bevat boodschappen die angst oproepen door middel van sociale dreiging (een ander verwonden door onder invloed te rijden, een ander in aanmerking laten komen voor een andersvalide parkeerplaats door dronken rijgedrag, en sociale uitsluiting ten gevolge van reputatieschade door binge-drinken). Deze boodschappen hebben als doelstelling de respondenten te

doen schrikken en de gevolgen van bingedrinken op onze sociale omgeving te benadrukken. De derde conditie bevat boodschappen die een humoristisch effect beogen met verwijzing naar een fysieke dreiging (haarverlies, rimpels en obesitas bij vrouwen; de dood ten gevolge van alcoholvergiftiging; en niet meer in staat zijn om zelf thuis te geraken). Zo probeert men op een humoristische manier de fysieke gevolgen van bingedrinken te benadrukken. De vierde conditie bevat boodschappen die een humoristisch effect beogen door middel van een sociale bedreiging (in slaap vallen op feestjes en daardoor jezelf belachelijk maken, overgeven in een publieke ruimte en ten laatste de aantrekkingskracht van de andere na een aantal glazen). Hierbij probeert men op een grappige manier de sociale gevolgen van bingedrinken voor te stellen. De boodschappen en beelden werden overgenomen uit bestaande voorlichtingscampagnes.

Resultaten

Manipulatiechecks. Er werden twee manipulatiechecks uitgevoerd om na te gaan of de gekozen advertenties voldoen aan de vooropgestelde voorwaarden van de condities.

De eerste manipulatiecheck gaat na of de voorlichtingsadvertenties in de humorconditie als meest humoristisch werden bevonden en of de mate van opgeroepen angst het hoogst is bij de advertenties in de angstcondities. Uit de eerste controle blijkt dat er gemiddeld meer angst wordt opgeroepen in de angstcondities ($M = 4.31$, $SD = .91$) dan in de humorcondities ($M = 3.38$, $SD = 1.09$; $t(157) = 5.92$, $p < .001$). Ook worden de advertenties in de humorcondities ($M = 4.03$, $SD = 1.07$) gemiddeld als meer humoristisch ervaren dan in de angstcondities ($M = 2.68$, $SD = .68$; $t(138) = -9.61$, $p < .001$). Uit deze resultaten kunnen we dus besluiten dat de eerste manipulatiecheck 'angst versus humor' geslaagd is.

De tweede manipulatiecheck gaat na of de advertenties met een fysieke dreiging als een fysieke dreiging worden beschouwd en of de mate van gepercipieerde sociale dreiging het grootst is in de socialedreigingcondities. Bij de tweede manipulatiecheck bleek dat de advertenties in de fysieke condities ($M = 4.95$, $SD = 1.16$) gemiddeld meer als fysiek bedreigend worden beschouwd dan in de sociale condities ($M = 3.90$, $SD = 1.32$; $t(162) = 5.37$, $p < .001$). Wanneer men kijkt naar de mate van gepercipieerde sociale dreiging ziet men dat de advertenties in de socialedreigingcondities ($M = 4.18$, $SD = 1.35$) gemiddeld meer als sociaal bedreigend worden ervaren, vergeleken met de fysiekedreigingcondities ($M = 3.95$, $SD = 1.34$). Het verschil tussen beide condities volgt dus de verwachte richting, maar is niet significant,

$t(162) = -1.12, p = .27$. De manipulatie ‘fysiek versus sociaal’ bleek dus niet volledig geslaagd te zijn.

Antwoorden op onderzoeksvragen. De eerste onderzoeksvraag gaat na of humoristische of angstaanjagende boodschappen aanleiding geven tot een positievere attitude ten opzichte van de voorlichtingsadvertentie. Om dit te meten werd gebruikgemaakt van een schaal opgesteld door Zhang (1996). Deze 4 item-schaal (niet aangenaam – aangenaam, niet leuk – leuk, niet interessant – interessant, irriterend – niet irriterend) werd gemeten aan de hand van een 7-punten Likertschaal. Na een betrouwbaarheidsanalyse ($\alpha = .74$) werd ‘attitude ten opzichte van de advertentie’ gevormd door het gemiddelde te nemen van deze vier items. Humoristische advertenties ($M = 3.85, SD = .87$) brengen een positievere attitude teweeg ten opzichte van de advertentie dan angstaanjagende advertenties ($M = 3.22, SD = .60; t(144) = -5.35, p < .001$).

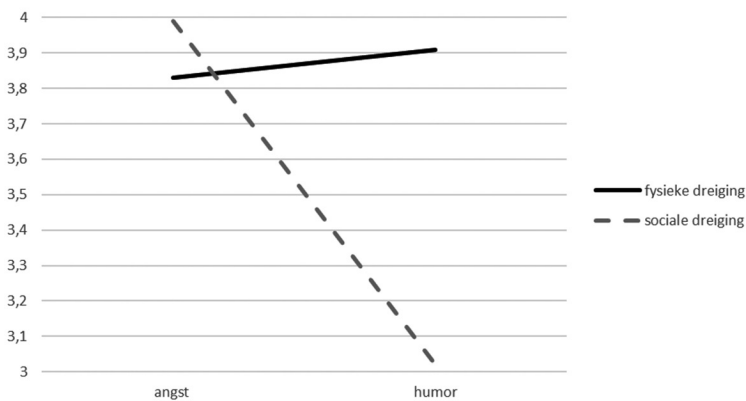
De tweede onderzoeksvraag bekijkt welke dreiging (fysiek of sociaal) de grootste overtuigingskracht heeft bij jongvolwassenen. Om te achterhalen of de informatie uit de getoonde advertenties aannemelijk, invloedrijk en overtuigend waren, maakte men gebruik van een schaal opgesteld door Gurhan-Canli & Maheswaran (2000). De respondenten werden gevraagd volgende uitspraak op een 7-punten semantisch differentiaal te beoordelen ‘De informatie/argumenten in de boodschap vond u: zeer zwak – zeer sterk, niet aannemelijk – zeer aannemelijk, niet invloedrijk – zeer invloedrijk, niet overtuigend – zeer overtuigend’. De schaal was intern consistent ($\alpha = .88$). Vervolgens maakte men deze variabele aan door het gemiddelde te nemen van de vier items. De respondenten bleken gemiddeld meer overtuigd te zijn door de advertenties met een fysieke dreiging ($M = 3.87, SD = 1.21$) dan de advertenties met een sociale dreiging ($M = 3.51, SD = 1.04; t(162) = 2.03, p < .05$).

De derde onderzoeksvraag gaat na welke appeals (humor of angst) de grootste overtuigingskracht hebben. De werkwijze is dezelfde als in de paragraaf hierboven geschetst. De aangeboden informatie in de advertentie is gemiddeld meer overtuigend in de angstcondities ($M = 4.82, SD = .90$) dan in de humorcondities ($M = 4.13, SD = 1.06; t(162) = 4.51, p < .001$).

De vierde vraag onderzoekt of humoristische of angstaanjagende boodschappen tot een grotere verwachte effectiviteit leiden. Voor de verwachte effectiviteit van de advertentie moesten de respondenten aan de hand van een 7-punten Likertschaal vier vragen beantwoorden van 1 (helemaal niet) tot en met 7 (heel veel) zoals “In welke mate zorgt deze advertentie ervoor dat u uw gedrag wilt aanpassen?” (Reichert, Heckler, & Jackson, 2001). Na de betrouwbaarheidsanalyse met een hoge chronbach's alpha ($\alpha = .91$) werd het gemiddelde genomen van de vier items om de afhankelijke variabele ‘verwachte

effectiviteit' te creëren. De respondenten bleken na het zien van angstaanjagende anti-bingedrinkingadvertenties ($M = 3.91$, $SD = .96$) significant meer overtuigd over de negatieve gevolgen van bingedrinken en gaven aan meer bereid te zijn om actie te ondernemen dan na het zien van humoristische anti-bingedrinkingadvertenties ($M = 3.47$, $SD = 1.26$; $t(151) = 2.54$, $p < .05$).

Met de vijfde onderzoeksvraag wordt ingeschat welke soort boodschap (humor of angst) in combinatie met welke dreiging (fysiek of sociaal) leidt tot de grootste verwachte effectiviteit. Uit de analyse blijkt dat 'appeal' en 'dreiging' interageren met elkaar ($F(1,160) = 9.86$, $p < .01$). Zo is het verschil tussen de humoristische en angstaanjagende advertenties anders naargelang de soort dreiging (fysiek of sociaal) waaraan men wordt blootgesteld. Uit posthoc tests bleek de conditie angst met sociale dreiging ($M = 3.99$, $SD = .78$) enkel significant te verschillen van de conditie humor met sociale dreiging ($M = 3.02$, $SD = 1.05$; $t(74) = 4.76$, $p < .001$). Dit betekent dat de conditie angst met sociale dreiging niet significant van de conditie angst met fysieke dreiging verschilt ($M = 3.83$, $SD = 1.12$; $t(69) = .78$, $p = .44$). Gelijkerwijze verschillen de angstaanjagende advertenties met een sociale dreiging niet significant ten opzichte van de humoristische advertenties met een fysieke dreiging ($M = 3.91$, $SD = 1.30$, $t(65) = .34$, $p = .74$). De humoristische advertenties met een fysieke dreiging verschillen niet significant van de angstaanjagende advertenties met een fysieke dreiging ($t(73) = .29$, $p = .77$). Humoristische advertenties met een sociale dreiging daarentegen verschillen significant ten opzichte angstaanjagende advertenties met een fysieke dreiging ($t(79) = -3.33$, $p < .01$). Eveneens verschillen humoristische



Figuur 1 Interactie-effect op effectiviteit van angst en humor appeals en fysieke en sociale dreigingen

advertenties met een sociale dreiging significant ten opzichte van de conditie humor met fysieke dreiging ($t(80) = -3.41, p < .01$). Kort samengevat, kunnen we stellen dat angstaanjagende advertenties met een sociale dreiging het meest effectief zijn. Humoristische advertenties met een sociale dreiging zijn het minst effectief om jongvolwassenen te overtuigen over de risico's van bingedrinken. Zie Figuur 1.

Discussie

Humor versus angst. Dit onderzoek bevestigt dat humoristische advertenties wat betreft attitude positiever werden beoordeeld dan de angstaanjagende advertenties. Door de opgewekte positieve emoties verkiezen verschillende sociale marketeers humor appeals boven fear appeals, omdat zij ervan uit gaan dat deze positieve gevoelens mogelijke ongewenste reacties zouden vermijden. Bijgevolg zou een positieve attitude ook kunnen leiden tot een gunstige intentie tot het stellen van het gewenste gedrag (Lee & Ferguson, 2002). Aangezien het empirisch bewijs omtrent het gebruik van humor bij bedreigende onderwerpen vrij schaars is en er nog altijd onenigheid bestaat over de werking en effectiviteit van fear appeals, bieden onze resultaten een interessante bijdrage aan deze aanhoudende discussie (Kohn et al., 1982; Stainback & Rogers, 1983; Yoon, 2015).

Wat betreft overtuigingskracht en effectiviteit, zien we een totaal ander beeld dan bij attitude. Angstaanjagende advertenties bleken niet alleen over meer overtuigingskracht te beschikken, maar konden de jongvolwassenen ook meer overtuigen over de negatieve gevolgen (verwachte effectiviteit) van bingedrinken dan de humoristische advertenties.

Fysieke versus sociale gevolgen. We testten in deze studie ook of fysieke of sociale dreigingen het meest effectief zijn om de intentie tot excessief alcoholgebruik bij studenten te reduceren. Verschillende onderzoekers gingen ervan uit dat fysieke dreigingen doeltreffender zijn dan sociale dreigingen, omdat men bij fysieke dreigingen wordt geconfronteerd met verwondingen en de dood (Reardon & Miller, 2008). Andere studies vonden dat sociale dreigingen voornamelijk effectief waren bij jongeren, aangezien men in deze levensfase gevoelig is voor sociale aanvaarding en sociale uitsluiting (Schoenbachler & Whittler, 1996; Smith & Stutts, 2003). Uit onze resultaten kunnen we besluiten dat fysieke dreigingen in anti-bingedrinkcampagnes beter werken bij jongvolwassenen dan sociale dreigingen.

Interactie humor/angst en fysieke/sociale gevaren. Voorgaande onderzoeksvragen geven aanleiding tot de kernvraag van deze studie, namelijk

welke soort boodschap (angst of humor) in combinatie met welke dreiging (fysiek of sociaal) leidt tot de grootste effectiviteit. De analyse toonde aan dat de variabelen 'appeal' en 'dreiging' met elkaar interageren. Dit betekent dat de mate van gepercipieerde effectiviteit tussen de angstaanjagende en humoristische advertenties verschillend is naargelang de soort dreiging.

De resultaten wijzen uit dat humoristische advertenties met een sociale dreiging het minst effectief zijn om studenten bewust te maken over de risico's verbonden aan bingedrinken. Een verklaring hiervoor kan zijn dat de humor de gepercipieerde ernst van de anti-bingedrinkingboodschap doet afnemen, waardoor die zijn doel mist (Yoon, 2015). Humor zorgt er immers voor dat de sociale dreiging niet meer als een echte dreiging wordt gepercipieerd.

Angstaanjagende advertenties met een sociale dreiging daarentegen blijken gemiddeld gezien het beste te werken bij jongvolwassenen, maar de effectiviteit verschilt niet significant ten opzichte van angstaanjagende advertenties met een fysieke dreiging en humoristische advertenties met een fysieke dreiging.

Het is van belang te beseffen dat deze resultaten niet zonder meer veralgemeend kunnen worden, aangezien ze enkel van toepassing zijn op de bingedrinkingproblematiek binnen de leeftijdscategorie van 18- tot 26-jarigen en opgewekt werden door de specifieke campagnes gebruikt in dit onderzoek.

Conclusie

Het bingedrinkgedrag binnen studentenkringen is een heikel probleem in onze samenleving. Daarom is het cruciaal dat studenten in de toekomst verder geïnformeerd worden over de nefaste gevolgen van overmatig alcoholgebruik. Zo brengen zij immers niet alleen hun eigen welzijn in gevaar (fysieke dreiging), maar kunnen zij ook anderen in gevaar brengen (sociale dreiging).

Deze studie differentieert zich ten opzichte van de voorgaande literatuur op drie vlakken. Ten eerste werden angst en humor appeals in het verleden vooral apart onderzocht. Ondanks de vele studies naar het gebruik van fear appeals in socialemarketingboodschappen en het gebruik van humor als één van de meest gehanteerde communicatiestrategieën in de reclamewereld, werden beide nog maar zeer weinig met elkaar vergeleken. Ten tweede heeft men de effectiviteit van deze appeals nog niet bestudeerd in combinatie met de soort dreiging (fysiek en sociaal). Ten derde zijn deze

socialemarketingstrategieën nog maar weinig onderzocht wat betreft co-mazuipen bij studenten.

We kunnen besluiten dat humoristische advertenties positieve gevoelens opwekken. Ondanks deze positieve attitude ten opzichte van de humoristische advertenties, blijken de angstaanjagende advertenties over meer overtuigingskracht te beschikken en zijn ze ook het meest effectief. Fysieke dreigingen zijn effectiever dan sociale dreigingen. Wanneer men appeal en dreiging met elkaar combineert, dan zien we dat humoristische advertenties met een sociale dreiging het minst effectief zijn.

Deze studie kan een meerwaarde betekenen binnen het onderzoek naar gezondheidscommunicatie. Op basis van onze bevindingen kunnen we sociale marketeers aanraden om bij een dergelijke problematiek angstaanjagende advertenties te gebruiken. Een sociale dreiging gecombineerd met humor lijkt de ernst te minimaliseren en daardoor de overtuigingskracht en effectiviteit van de voorlichtingsadvertentie te doen afnemen. Als men humor wil gebruiken om meer goodwill (positieve attitude) te creëren, kunnen we stellen dat humor appeals met fysieke dreiging een goed alternatief zijn.

Bibliografie

- Belch, G.E., & Belch, M.A. (1984). An investigation of the effects of repetition on cognitive and affective reactions to humorous and serious television commercials. *Advances in Consumer Research*, 1(1), 4-10.
- Bloom, B.L. (1988). *Health psychology: A psychological perspective*. Englewood Cliffs, NJ: Prentice Hall.
- Boles, S.M., & Miotto, K. (2003). Substance abuse and violence: A review of the literature. *Aggression & Violent Behavior*, 8(2), 155-174.
- Brennan, L., & Binney, W. (2010). Fear, guilt, and shame appeals in social marketing. *Journal of Business Research*, 63(2), 140-146.
- Burnett, J.J., & Oliver, R.L. (1979). Fear appeal effects in the field: A segmentation approach. *Journal of Marketing Research*, 16(2), 181-190.
- Caubergh, V., De Pelsmacker, P., Janssens, W., & Dens, N. (2009). Fear, threat and efficacy in threat appeals: Message involvement as a key mediator to message acceptance. *Accident Analysis and Prevention*, 41(2), 276-285.
- Easterling, D.V., & Leventhal, H. (1989). Contribution of concrete cognition to emotion: Neutral symptoms as elicitors of worry about cancer. *Journal of Applied Psychology*, 74(5), 787-796.
- Eisend, M. (2009). A meta-analysis of humor in advertising. *Journal of the Academy of Marketing Science*, 37(2), 191-203.
- Gurhan-Canli, Z., & Maheswaran, D. (2000). Determinants of country-of-origin evaluations. *Journal of Consumer Research*, 27(1), pp. 96-108.
- Hastings, G., Stead, M., & Webb, J. (2004). Fear appeals in social marketing: Strategic and ethical reasons for concern. *Psychology and Marketing*, 21(11), 961-986.

- Hendriks, H. & Janssen, L. (2018). Frightfully funny: Combining threat and humour in health messages for men and women. *Psychology & Health*, 33 (5), 594-613.
- Hunter, Z., & Francescutti, L.H. (2013). Facing the consequences of binge drinking. *Canadian Family Physician / Médecin de famille canadien*, 59(10), 1041-1042.
- Kohn, P.M., Goodstadt, M.S., Cook, G.M., Sheppard, M., & Chan, G. (1982). Ineffectiveness of threat appeals about drinking and driving. *Accident Analysis and Prevention*, 14(6), 457-464.
- Krishnan, H.S. & Chakravarti, D. (2003). A process analysis of the effects of humorous advertising executions on brand claims memory. *Journal of Consumer Psychology*, 13(3), 230-245.
- Lagerwerf, L. (2007). Irony and sarcasm in advertisements: effects of relevant inappropriateness. *Journal of Pragmatics*, 39(10), 1702-1721.
- Laroche, M., Toffoli, R., Zhang, Q., & Pons, F. (2001). A cross-cultural study of the persuasive effect of fear appeal messages in cigarette advertising: China and Canada. *International Journal of Advertising*, 20(3), 237-317.
- LaTour, M.S., & Zahra, S.A. (1989). Fear appeals as advertising strategy: Should they be used? *Journal of Consumer Marketing*, 6(2), 61-70.
- Lee, M.J., & Ferguson, M.A. (2002). Effects of anti-tobacco advertisements based on risk-taking tendencies: Realistic fear versus vulgar humor. *Journalism & Mass Communication Quarterly*, 79(4), 945-963.
- Leventhal, H. (1971). Fear appeals and persuasion: The differentiation of a motivational construct. *American Journal of Public Health*, 61(6), 1208-1224.
- Lewis, I.M., Watson, B., White, K.M., Richard, T. (2007). Promoting Public Health Messages: Should we move beyond fear-evoking appeals in road safety? *Qualitative Health Research*, 17(1), 61-74.
- Martin, I., & Kamins, M. (2010). An application of terror management theory in the design of social and health-related anti-smoking appeals. *Journal of Consumer Behaviour*, 9(3), 172-190.
- Morreale, R. (2016). *De ontnuchterende voorlichtingscampagnes omtrent de bingedrinkingsproblematiek bij studenten: Een experimenteel onderzoek*. (ongepubliceerde masterproef). KU Leuven, België.
- Murgraff, V., Parrott, A., & Bennett, P. (1999). Risky single-occasion drinking amongst Young people – definition, correlates, policy and intervention: A broad overview of research findings. *Alcohol and Alcoholism*, 34(1), 3-14.
- Naimi, T.S., Nelson, D.E., & Brewer, R.D. (2010). The intensity of binge alcohol consumption among U.S. adults. *American Journal of Preventive Medicine*, 38(2), 201-207.
- Rayner, E., & Baxter, S.M., & Illicic, J. (2015). Smoker's recall of fear appeal imagery: Examining the effect of fear intensity and fear type. *Australasian Marketing Journal*, 23(1), 61-66.
- Reardon, J., & Miller, C. (2008). Smoking prevention messages for adolescents: How intensity, valence, and recipient of consequences affect attitude toward the ad and intent to smoke. *The Journal of Marketing Theory and Practice*, 16(1), 67-77.
- Reichert, T., Heckler, S., & Jackson, S. (2001). The effects of sexual social marketing appeals on cognitive processing and persuasion. *Journal of Advertising*, 30(1), pp. 13-27.
- Schoenbachler, D.D., & Whittler, T.E. (1996). Adolescent processing of social and physical threat communications. *Journal of Advertising*, 25(4), 37-55.
- Smith, K.H., & Stutts, M.A. (2003). Effects of short-term cosmetic versus long-term health fear appeals in anti-smoking advertisements on the smoking behavior of adolescents. *Journal of Consumer Behavior*, 3(2), 157-177.
- Speck, P.S. (1987). *On Humour and Humour in Advertising* [Unpublished doctoral dissertation]. Texas Tech University.
- Stainback, R.D., & Rogers, R.W. (1983). Identifying effective components of alcohol abuse prevention programs: Effects of fear appeals, message style, and source expertise. *International Journal of the Addictions*, 18, 393-405.

- Stephens, D.N., & Duka, T. (2008). Cognitive and emotional consequences of binge drinking: Role of amygdala and prefrontal cortex. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 363(1507), 3169-3179.
- Wechsler, H., Davenport, A., Dowdall, G., Hoeykens, B., & Castillo, S. (1994). Health and behavioral consequences of binge drinking in college campuses: A national survey of students at 140 campuses. *Journal of the American Medical Association*, 272(21), 1672-1677.
- Weinberger, M., & Gulas, C. (1992). The impact of humor in advertising: A review. *Journal of Advertising*, 21(4), 35-59.
- Witte, K. (1994). Fear control and danger control: A test of the extended parallel process model (EPPM). *Communication Monographs*, 61(2), 113-134.
- Wolburg, J.M. (2001). The 'risky business' of binge drinking among college students: Using risk models for PSAs and anti-drinking campaigns. *Journal of Advertising*, 30(4), 23-39.
- Yoon, H.J. (2015a). Emotional and Cognitive Responses to Non-Humorous and Humorous Threat Persuasion Advertisements. *Journal of Current Issues & Research in Advertising*, 36(1), 52- 69.
- Yoon, H.J. (2015b). Humor effects in shame-inducing health issue advertising: The moderating effects of fear of negative evaluation. *Journal of Advertising*, 44(2), 126.
- Yoon, H.J., & Tinkham, S. (2013). Humorous Threat Persuasion in Advertising: The Effects of Humor, Threat Intensity, and Issue Involvement. *Journal of Advertising*, 42(1), 30-41.
- Zhang, Y. (1996). The effect of humor in advertising: An individual-difference perspective. *Psychology and Marketing*, 13(6), pp. 531-545.

Effecten van anti-rookboodschappen bij jongeren

Een conceptueel replicatieonderzoek

Heleen van Wijk en Carel Jansen

Rijksuniversiteit Groningen/Universiteit Stellenbosch

TVT 41 (1): 323–336

DOI: 10.5117/TVT2019.1.022.WIJK

Abstract

Impact of anti-smoking message on youth: A conceptual replication study

Based on an experiment with young Dutch smokers, Mollen, Engelen, Kessels, and Van den Putte (2017) advise the European Commission to reconsider the use of warning labels on cigarette packages. Participants in their study were students at institutes for higher education ($N = 132$). Their attitudes and behavioral intentions were found to be affected both by message framing and temporal discounting in the verbal anti-smoking messages they were presented with. Gain frames were more persuasive than loss frames, and short-term consequences were more persuasive than long-term consequences.

We performed a conceptual replication study in which we employed the same experimental design, followed the same procedure, and used similar anti-smoking messages. Other than in Mollen et al., however, both smokers ($N = 129$) and non-smokers ($N = 173$) participated, and the verbal messages were accompanied with matching visuals. Furthermore, participants were younger ($M = 18.0$ years) than in the original study ($M = 22.4$ years), and they were students at a lower educational level. We found none of the effects reported in the original study, neither for smokers nor for non-smokers. Therefore, our study does not support the advice from Mollen et al. (2017).

Keywords: anti-smoking messages, smokers, non-smokers, message framing, temporal context, replication

Inleiding

Wie in Nederland een pakje sigaretten koopt, wordt op de verpakking uitdrukkelijk gewaarschuwd voor de gevolgen van roken. Vroeger gebeurde dat alleen met onheilspellende teksten; sinds mei 2016 worden daar angst-aanjagende plaatjes aan toegevoegd. Volgens recent onderzoek hebben de waarschuwingen effect. In onderzoek van KANTAR TNS (2017) zei 46% van de ondervraagden en 32% van de rokers over alleen de waarschuwendende teksten dat die hen bewust maakten van de schadelijkheid van roken; bij de teksten plus plaatjes was dat respectievelijk 51% en 38% (p. 30). In onderzoek van het Nationaal Expertisecentrum Tabaksontmoediging (2017) zei 14% van de ex-rokers dat de gezondheidswaarschuwingen op verpakkingen en in mediacampagnes ertoe hadden bijgedragen dat ze gestopt waren; van de rokers verwachtte 21% dat die waarschuwingen hen mede zouden motiveren om te stoppen (z.p., Figuur 11).

Anti-rookboodschappen op sigarettenpakjes doen er dus toe, en onderzoek naar condities waaronder ze meer of minder effectief zijn, is de moeite waard. Over zulk onderzoek, en wel bij jonge, hoogopgeleide rokers in Nederland, wordt gerapporteerd door Mollen, Engelen, Kessels en Van den Putte (2017). Hun experiment, waarin attitudes en intenties de belangrijkste afhankelijke variabelen waren, wees uit dat (tekstuele) boodschappen waarin kortetermijneffecten werden benadrukt effectiever waren dan boodschappen waarin gewezen werd op langtermijneffecten. Ook bleken *gain-framed* boodschappen, over de positieve effecten van stoppen met roken, effectiever dan *loss-framed* boodschappen, over de negatieve effecten van roken. Het meest overtuigend bleken gain-framed boodschappen over kortetermijneffecten. Mollen et al. (2017) adviseren de Europese Commissie om in elk geval jonge, hoogopgeleide jongeren niet langer te wijzen op gezondheidsproblemen die ze op lange termijn kunnen verwachten van hun rookgedrag, maar op positieve kortetermijneffecten van stoppen met roken (p. 26).

Mede gezien de praktische impact van dit advies besloten wij tot een conceptueel replicatieonderzoek (Eerland & Van den Bergh, 2016). Een tweede aanleiding lag in het contrast tussen de uitkomsten van Mollen et al. (2017) en de resultaten van eerder onderzoek. Zo is de bevinding dat de onderzochte gain-framed boodschappen effectiever waren dan de loss-framed alternatieven opmerkelijk in het licht van ander onderzoek naar de effecten van *message framing* dat vrijwel geen verschillen liet zien. Uit meta-analyses van onderzoek naar winst- versus verliesframing bij persuasieve gezondheidsboodschappen concludeert O’Keefe (2016) dat “gain-framed

and loss-framed appeals do not generally differ in persuasiveness" (p. 225). Dat geldt zowel voor boodschappen die zijn gericht op detectiegedrag (zoals kankerscreening) als voor boodschappen die zijn gericht op preventiegedrag (zoals stoppen met roken), aldus O'Keefe met verwijzing naar Gallagher en Updegraff (2012), O'Keefe en Jensen (2007) en O'Keefe en Jensen (2009). Bij geen van beide soorten gezondheidsboodschappen worden de theoretische verwachtingen ('bij detectiegedrag is loss-framing effectiever; bij preventiegedrag gain-framing') ondersteund door het beschikbare empirisch onderzoek.

Het onderzoek van Mollen et al. (2017)

Verwijzend naar onder meer de meta-analyse van Gallagher en Updegraff (2012), die volgens Mollen et al. (2017) wel een positief effect vonden van gain-framed versus loss-framed gezondheidsboodschappen, verwachtten Mollen et al. dat in hun onderzoek gain-framed boodschappen overtuigender zouden zijn dan loss-framed tegenhangers (pp. 20-21). Gallagher en Updegraff inventariseerden behalve de effecten op attitudes en intenties ook de effecten op daadwerkelijk gedrag dat volgde in de maanden nadat de respondenten blootgesteld waren aan het voorlichtingsmateriaal. Voor attitudes of gedragsintenties vonden Gallagher en Updegraff geen significante effectverschillen, net zoals het geval was bij O'Keefe en Jensen (2007; 2009).

Alleen voor bij het feitelijke preventiegedrag van de respondenten vonden Gallagher en Updegraff (2012) een voordeel van gain-framed gezondheidsboodschappen; dat voordeel was echter bescheiden. De auteurs stellen ook dat "[...] an insufficient number of published studies were available within particular domains to draw firm conclusions" (p. 111).

Terwijl de meeste anti-rookboodschappen betrekking hebben op schadelijke langetermijneffecten, verwachtten Mollen et al. (2017) dat refereren aan kortetermijneffecten beter zou werken. Onder verwijzing naar onder meer Green en Myerson (2004) noemen ze dit verschijnsel van *temporal discounting* logisch vanuit een economisch en evolutionair perspectief: mensen zijn geneigd een onmiddellijke beloning te prefereren boven een uitgestelde beloning (Mollen et al., 2017, p. 21). Ook verwachtten ze een interactie-effect van *message framing* en *temporal context* (verder: tijdframe). Het positieve effect van gain-framed boodschappen zou het grootst zijn bij verwijzingen naar kortetermijneffecten (p. 32).

Eerder onderzoek naar het effect van tijdframe bij jongeren wijst echter niet in deze richting. Zo bestudeerden Zhao en Peterson (2007) anti-rookwaarschuwingen bij rokende jongeren in de VS van gemiddeld 26 jaar. Ze vonden dat tijdframe de gedragsintenties niet of nauwelijks beïnvloedde. Burgers en Veldhuis (2013) deden onderzoek bij Nederlandse scholieren van gemiddeld 15 jaar (rokers en niet-rokers). Ook zij vonden ook geen effect van tijdframe. Smith en Stutts (2003) deden onderzoek bij jongeren (rokers en niet-rokers) in de VS van ongeveer 15 jaar naar reacties op waarschuwingen waarin gewezen werd op ofwel cosmetische kortetermijneffecten ofwel gezondheidsgerelateerde langetermijneffecten van roken. De laatste soort waarschuwingen waren het meest effectief (p. 172). Onduidelijk is echter in hoeverre dat veroorzaakt werd door het contrast tussen korte- en langetermijneffecten of door het contrast tussen cosmetische en gezondheidsgerelateerde effecten.

Mollen et al. (2017) voerden een experiment uit met een 2 (gain of loss) $\times 2$ (korte termijn of lange termijn) tussenproefpersoonsondesign. Alle 132 deelnemers waren rokende hoger-onderwijsstudenten (47.7% mannen; gemiddeld 22.4 jaar; $SD = 1.85$). Op verzoek van de proefleider kozen ze twaalf keer uit twee verschillende pakjes het merk dat hun voorkeur had. De deelnemers wisten niet dat het de onderzoekers daar feitelijk niet om ging, maar dat het onderzoek gericht was op de effecten van het type anti-rookboodschap dat ze twaalf keer gezien hadden. Afhankelijk van de experimentele conditie waren die boodschappen steeds ofwel gain-framed ofwel loss-framed, en waren ze steeds ofwel op de korte termijn ofwel op de lange termijn gericht. Thema's van de kortetermijnboodschappen waren geldverlies (loss) of kostenbesparing (gain), en een verslechterend (loss) of een verbeterend (gain) uithoudingsvermogen. De langetermijnboodschappen gingen over ongezonde (loss) of gezonde (gain) tanden, en over ongezonde (loss) of schone (gain) longen.

Over de statistische power van de uitgevoerde variantieanalyses geven Mollen et al. (2017) geen informatie. Die power ($1-\beta$) is met de gegevens in het artikel echter wel te berekenen.¹ Gegeven de omvang van de steekproef ($N = 132$), α van .05, en het ontbreken van covariaten is $1-\beta$ bij een groot effect ($f = .40$) gelijk aan .99, bij een middelgroot effect ($f = .25$) .81, en bij een klein effect ($f = .10$) .21.

De manipulaties bleken te worden herkend: wanneer alleen gain-framed boodschappen werden aangeboden was de gemiddelde score voor *valence* (hoe negatief of positief zijn de gevolgen van stoppen met roken?) significant hoger dan wanneer de deelnemers alleen loss-framed boodschappen te zien kregen. Hetzelfde gold mutatis mutandis voor de condities waarin

op kortetermijneffecten werd gewezen, vergeleken met de condities waarin aan langetermijneffecten werd gerefereerd.

Zowel voor attitude als gedragsintentie werden significante hoofdeffecten gevonden. Deelnemers die gain-framed boodschappen hadden gezien, stonden positiever tegenover stoppen met roken dan deelnemers aan wie loss-framed boodschappen waren getoond, en hetzelfde gold voor hun intentie om te stoppen. Vergelijkbare uitkomsten werden gevonden voor tijdframe: kortetermijnboodschappen leidden tot een positievere attitude en intentie dan langetermijnboodschappen. Een interactie-effect van message framing en tijdframe werd niet gevonden voor attitude, maar wel voor intentie. Het meest effectief in dit opzicht bleken zoals verwacht gain-framed boodschappen waarin aan kortetermijneffecten werd gerefereerd (pp. 24-25). Zie Tabel 1.

Tabel 1 Gemiddelden en standaarddeviaties bij rokers (N = 132) voor attitude en intentie in het oorspronkelijke onderzoek (Mollen et al., 2017, p. 24)

Conditie	Attitude tegenover stoppen met roken (min. 1, max. 7)	Intentie om te stoppen met roken (min. 1, max. 7)
Gain-framed, kortetermijneffect	$M = 6.04; SD = 0.94$	$M = 5.53; SD = 1.57$
Gain-framed, langetermijneffect	$M = 5.51; SD = 1.34$	$M = 3.01; SD = 1.42$
Loss-framed, kortetermijneffect	$M = 5.65; SD = 1.28$	$M = 3.51; SD = 1.89$
Loss-framed, langetermijneffect	$M = 4.50; SD = 2.24$	$M = 2.21; SD = 1.63$

Uit deze resultaten concluderen Mollen et al. (2017) dat de Europese Commissie het gebruik van waarschuwingen over negatieve langetermijngevolgen van roken zou moeten herzien. Hun resultaten impliceren immers dat het in ieder geval bij jonge, hoogopgeleide rokers effectiever lijkt om kortetermijnwinsten van stoppen met roken onder de aandacht te brengen (p. 26).

Het replicatieonderzoek: methode

In het replicatieonderzoek volgden we dezelfde aanpak als die van Mollen et al. (2017): alle deelnemers kregen twaalf paren sigarettenpakjes te zien met daarop steeds dezelfde anti-rookboodschap. Anders dan Mollen et al. voegden wij aan de tekstuele anti-rookboodschappen visualisaties toe; zie Figuur 1. Zulke visualisaties zijn tegenwoordig een verplicht onderdeel van de waarschuwingen op tabaksverpakkingen. Een recente meta-analyse leidt tot de conclusie dat visuele anti-rookwaarschuwingen ook effectiever zijn dan alleen verbale waarschuwingen (Noar et al., 2016).



Figuur 1 Voorbeelden van onderzoeksmateriaal gebruikt in het replicatieonderzoek: een gain-framed boodschap over een kortetermijneffect op twee sigarettenpakjes, en daaronder een loss-framed boodschap over een langetermijneffect op twee sigarettenpakjes

Afhankelijk van de experimentele conditie kreeg ook in ons onderzoek elke deelnemer twaalf keer twee sigarettenpakjes te zien, met daarop steeds hetzelfde type boodschap: ofwel op de korte- ofwel op de lange termijn gericht, en ofwel gain-framed ofwel loss framed. Net als Mollen et al. (2017) hielden wij verborgen dat het om de effecten van de anti-rookboodschap ging. We zeiden dat we wilden weten naar welke pakjes de voorkeur uitging.

De thema's die aan de orde kwamen, waren grotendeels dezelfde als bij Mollen et al. (2017): geld, gebit en longen. Alleen waar in het oorspronkelijke onderzoek boodschappen gebruikt werden over effecten op het uithoudingsvermogen, werd daar in het replicatieonderzoek van afgeweken: zulke effecten bleken moeilijk te visualiseren. In plaats daarvan construeerden we boodschappen over effecten op de huid.

Bekend is dat rookgedrag samenhangt met opleidingsniveau: hoe lager de opleiding, hoe meer er wordt gerookt ("Volksgezondheidszorg.info", 2018). Daarom kozen wij als deelnemers voor leerlingen uit het middelbaar beroepsonderwijs (verder: mbo). Verder besloten we om behalve rokers ook niet-rokers in ons onderzoek te betrekken, dit omdat de overheid niet alleen probeert om rokers tot stoppen te brengen, maar ook om niet-rokers van het roken af te houden (zie bijvoorbeeld "Roken ontmoedigen door nieuwe regels", z.j.). Daarom werden de gain-framed waarschuwingsteksten die rokers en niet-rokers kregen voorgelegd iets algemener geformuleerd dan bij Mollen et al. (2017): niet 'Als je stopt met roken, bespaart dit je veel geld' bijvoorbeeld, maar 'Als je niet rookt, bespaart dit je veel geld'. De formulering van de loss-framed boodschappen bleef hetzelfde als in het oorspronkelijke onderzoek: 'Als je rookt, krijg je rotte tanden' bijvoorbeeld. Ook werd in het replicatieonderzoek de deelnemers niet gevraagd aan welk sigarettenmerk zij de voorkeur gaven. Uit een eerste pretest bleek dat niet-rokers dat een rare vraag vonden. Gevraagd werd daarom welke verpakking de deelnemers het mooist vonden. Wat ze daadwerkelijk van de vormgeving van de verpakking vonden was voor dit onderzoek niet relevant. Net als bij Mollen et al. ging het erom hoe de deelnemers reageerden op het type anti-rookboodschap dat ze twaalf keer te zien hadden gekregen.

Aan ons onderzoek namen 302 mbo-leerlingen deel (47.2% mannen; gemiddelde leeftijd 18.0 jaar; $SD = 1.98$). Zowel bij de rokers ($N = 129$) als bij de niet-rokers ($N = 173$) werden de vier versies van het materiaal (gain-framed kortetermijneffect; loss-framed kortetermijneffect, gain-framed langetermijneffect; loss-framed langetermijneffect) willekeurig verdeeld. Anders dan Mollen et al. (2017), die steeds zevenpuntsschalen gebruikten, lieten wij de items waarmee de afhankelijke variabelen werden gemeten, volgen

door vijfpuntsschalen – dit om het de deelnemers gemakkelijker te maken een duidelijke keuze tussen de alternatieven te maken (DeCastellarnau, 2018, pp. 1531-1532).

Bij de formulering van de items sloten we zo nauw mogelijk aan bij het oorspronkelijke onderzoek. Net als in Mollen et al. (2017) werd bij rokers attitude gemeten met deze vijf items: 'Ik vind stoppen met roken slecht .. goed; nadelig .. voordelig; ongezond .. gezond; onverstandig .. verstandig; onplezierig .. plezierig' ($M = 3.94$; $SD = 0.78$; $\alpha = .77$). Ook werden aan de rokers dezelfde vijf items voorgelegd om intentie te meten als die Mollen et al. gebruikten: 'Ik ben van plan [...] / Ik zal proberen [...] / Ik heb de intentie [...] / Ik ben bereid [...] / Ik verwacht [...] om binnen drie maanden te stoppen met roken' (schaal steeds 'volledig mee oneens .. volledig mee eens') ($M = 2.40$; $SD = 1.28$; $\alpha = .96$).

Ook bij de niet-rokers werden attitude en intentie gemeten, maar dan met betrekking tot *niet-roken*. Voor attitude waren de items: 'Ik vind roken slecht .. goed; nadelig .. voordelig; ongezond .. gezond; onverstandig .. verstandig; onplezierig .. plezierig' ($M = 1.63$; $SD = 0.74$; $\alpha = .86$). Ompoling leverde een gemiddelde attitude-score van 4.37 op. Intentie werd bij de niet-rokers gemeten met deze items: 'Ik ben van plan [...] / Ik zal proberen [...] / Ik heb de intentie [...] / Ik ben bereid [...] / Ik verwacht [...] om niet te gaan roken' (schaal: 'volledig mee oneens .. volledig mee eens') ($M = 4.34$; $SD = 1.00$; $\alpha = .90$).

Gemeten werd ook in hoeverre de deelnemers de boodschappen hadden gepercipieerd als loss-framed of gain-framed en als geldend voor de korte of de lange termijn. Net als Mollen et al. (2017) gebruikten we daarvoor deze twee items: 'De gevolgen die ik zag zijn negatief .. positief' en 'De gevolgen die ik zag zijn van toepassing op de korte termijn .. de lange termijn'.

Omdat de *consequentiebetrokkenheid* van de deelnemers hun reacties op de aangeboden waarschuwingen zou kunnen beïnvloeden (Hoeken, Hornikx, & Hustinx, 2012; Hertgers & Jansen, 2006; Cauberge, De Pelsmacker, Janssens, & Dens, 2009), besloten we ook te meten hoe belangrijk de deelnemers een correcte attitude tegenover (stoppen met) roken vonden. Daartoe werden aan de rokers vijf items voorgelegd, met formuleringen die nauw aansloten bij de adviezen van Hoeken et al. (pp. 220-221): 'Het is voor de kwaliteit van mijn leven van belang of ik rook of niet', 'Of ik wel of niet rook, beïnvloedt mijn dagelijks leven', 'Ik vind het gemakkelijk om te bedenken hoe roken mijn leven beïnvloedt', 'Dat ik rook, heeft invloed op mijn leven', en 'Mijn leven zou veranderen als ik zou stoppen met roken' ($M = 3.06$; $SD = 0.98$; $\alpha = .75$). De niet-rokers kregen deze items voorgelegd: 'Het is voor de kwaliteit van mijn leven van belang of ik rook of niet', 'Of ik wel of niet rook, beïnvloedt mijn dagelijks leven', 'Ik vind het gemakkelijk

om te bedenken hoe het feit dat ik niet rook mijn leven beïnvloedt', 'Dat ik niet rook, heeft invloed op mijn leven', en 'Mijn leven zou veranderen als ik zou beginnen met roken' ($M = 3.60$; $SD = 0.99$; $\alpha = .73$). Een commissie voor ethische toetsing gaf toestemming voor het onderzoek.

De statistische power ($1-\beta$) van de uitgevoerde variantieanalyses verschilde enigszins voor de rokers en de niet-rokers. Bij de rokers was de power gegeven de omvang van de steekproef ($N = 129$), α van .05 en één covariaat (consequentiebetrokkenheid) bij een groot effect ($f = .40$) gelijk aan .99, bij een middelgroot effect ($f = .25$) .80, en bij een klein effect ($f = .10$) .20. Bij de niet-rokers was de power gegeven de omvang van de steekproef ($N = 173$), α van .05 en het ontbreken van covariaten (zie hieronder) bij een groot effect ($f = .40$) gelijk aan .999, bij een middelgroot effect ($f = .25$) .90, en bij een klein effect ($f = .10$) .26.

Het replicatieonderzoek: resultaten

Manipulatiecheck

Om na te gaan of de gemanipuleerde message frames en korte- en langetermijneffecten als zodanig waren ervaren, werden twee ANOVA's uitgevoerd met de bedoelde perceptie als afhankelijke variabele. Framing bleek inderdaad tot perceptieverschillen te leiden: $F(1,300) = 16.54$; $p < .001$; $\eta^2 = .05$; bij de gain-framed boodschappen werden de gepresenteerde gevolgen als positiever gezien ($M = 2.24$; $SD = 1.47$) dan bij de loss-framed boodschappen ($M = 1.64$; $SD = 1.07$). De manipulatie van korte- versus langetermijneffecten bleek minder succesvol; het gepercipieerde verschil tussen beide soorten boodschappen was niet significant: $F(1,300) = 3.15$; $p = .08$; kortetermijneffecten: $M = 3.87$; $SD = 1.25$, langetermijneffecten: $M = 4.11$; $SD = 1.08$. Dat was verrassend, gegeven de resultaten van een pretest onder 40 andere mbo-leerlingen (rokers en niet-rokers). Deze leerlingen kregen dezelfde boodschappen met dezelfde instructies voorgelegd als de deelnemers aan het hoofdonderzoek. In de pretest leverden zowel de manipulatiecheck van de message framing als de manipulatiecheck van de korte- versus langetermijneffecten significante verschillen in de verwachte richting op.

Dat het hoofdonderzoek een verschil opleverde voor de perceptie van korte- versus langetermijneffecten dat slechts tendeerde naar significantie impliceert niet dat verdere analyses van de effecten van de manipulatie van tijdframe niet zinvol zouden zijn. Zoals O'Keefe (2016) laat zien, is ook als van elkaar verschillende boodschappen in een manipulatiecheck niet tot de beoogde respons leiden, de analyse van andere reacties op die

boodschappen relevant. Sterker nog: als verschillen tussen boodschappen worden gedefinieerd in termen van kenmerken die die boodschappen al dan niet hebben, zijn manipulatiechecks als zodanig onnodig (p. 187, noot 9). Wel beveelt O'Keefe (pp. 183-184) aan om zulke metingen te gebruiken om mogelijke mediatie-effecten te vinden tussen de variabele 'boodschap-kenmerk' en daardoor beïnvloede uitkomstvariabelen (zie ook O'Keefe, 2003; Tao & Bucy, 2007).

Vanwege van het replicatiekarakter van dit onderzoek worden hieronder eerst de effecten gepresenteerd op attitude en intentie van de rokers, de groep die het meest overeenkwam met de deelnemers aan het oorspronkelijke onderzoek. Separaat daarvan volgen de resultaten voor de groep niet-rokers.

Rokers

Bij de rokers werden twee ANCOVA's uitgevoerd waarin consequentiebetrokkenheid zoals beoogd als covariaat kon worden gebruikt: voldaan werd aan de voorwaarde dat de onafhankelijke variabelen geen significant effect hebben op de covariaat (Field, 2009, pp. 397-399): message framing $p = .72$; tijdframe $p = .14$. Voor attitude werden noch hoofd- noch interactie-effecten gevonden van message framing en van korte- versus langetermijneffecten: $F(1,124)$ resp. 0.13 ($p = .72$), 1.68 ($p = .20$) en 0.00 ($p = .96$). Wel was er een significant effect van consequentiebetrokkenheid op attitude: $F(1,124) = 9.60$; $p = .002$; $\eta^2 = .07$. Nadere analyse liet een positieve correlatie zien tussen deze twee variabelen: $r = .26$; $p = .003$.

Ook voor intentie werden geen hoofd- of interactie-effecten gevonden van message framing en van korte- versus langetermijneffecten: $F(1,124)$ resp. 1.18 ($p = .28$), 0.00 ($p = .99$) en $.05$ ($p = .82$). Wel was er ook nu een significant effect van consequentiebetrokkenheid: $F(1,124) = 3.95$; $p = .05$; $\eta^2 = .03$. Nadere analyse leverde een positieve correlatie op tussen consequentiebetrokkenheid en intentie: $r = .17$; $p = .05$. Zie Tabel 2.

Tabel 2 Gemiddelden en standaarddeviaties bij rokers (N = 129) voor attitude en intentie in het replicatieonderzoek

Conditie	Attitude tegenover stoppen met roken (min. 1, max. 5)	Intentie om te stoppen met roken (min. 1, max. 5)
Gain-framed, kortetermijneffect	$M = 3.90$; $SD = 0.93$	$M = 2.35$; $SD = 1.29$
Gain-framed, langetermijneffect	$M = 3.95$; $SD = 0.81$	$M = 2.25$; $SD = 1.15$
Loss-framed, kortetermijneffect	$M = 3.87$; $SD = 0.60$	$M = 2.55$; $SD = 1.37$
Loss-framed, langetermijneffect	$M = 4.06$; $SD = 0.72$	$M = 2.51$; $SD = 1.37$

Niet-rokers

Voor de niet-rokers werden twee ANOVA's uitgevoerd. Aan de voorwaarde voor het invoeren van een mogelijke covariaat (hier consequentiebetrokkenheid) dat de onafhankelijke variabelen daarop geen significante effecten hebben, werd nu niet voldaan: message framing $p = .01$; tijdframe $p = .001$. Noch voor attitude noch voor intentie werden hoofd- of interactie-effecten gevonden van message framing en korte- versus langetermijneffecten (attitude: $F(1,169)$ resp. 0.08 ($p = .77$), 0.12 ($p = .73$) en 0.06 ($p = .81$); intentie: $F(1,169)$ resp. 0.22 ($p = .64$), 0.19 ($p = .66$) en 1.97 ($p = .16$)). Zie Tabel 3.

Tabel 3 Gemiddelden en standaarddeviaties bij niet-rokers ($N = 173$) voor attitude en intentie in het replicatieonderzoek

Conditie	Attitude tegenover niet-roken (min. 1, max. 5)	Intentie om niet te gaan roken (min. 1, max. 5)
Gain-framed, kortetermijneffect	$M = 4.42; SD = 0.64$	$M = 4.44; SD = 0.91$
Gain-framed, langetermijneffect	$M = 4.35; SD = 0.70$	$M = 4.15; SD = 1.14$
Loss-framed, kortetermijneffect	$M = 4.36; SD = 0.74$	$M = 4.29; SD = 1.05$
Loss-framed, langetermijneffect	$M = 4.35; SD = 0.84$	$M = 4.44; SD = 0.91$

Mediatieanalyses

Conform de aanbevelingen van O'Keefe (2003; 2016) en Tao en Bucy (2007) gingen we ook na of mogelijke effecten van de manipulaties van message framing en tijdframe op attitude en intentie gemedieerd werden door de mate waarin de manipulaties als zodanig werden ervaren.² Zoals Hayes (2009) laat zien, kan er ook als er geen significant totaaleffect bestaat van een onafhankelijke variabele X op een afhankelijke variabele Y , sprake zijn van een indirect effect van X op Y via een mediërende variabele Y , bijvoorbeeld als diverse indirecte effecten in tegenovergestelde richting werken (pp. 413-414). Noch bij de rokers noch bij de niet-rokers vonden we significante directe, indirecte (via manipulatiepercepties) of totaal-effecten van waarschuwingsversie op attitude of intentie.

Discussie

Anders dan in het oorspronkelijke onderzoek van Mollen et al. (2017) werden in dit replicatieonderzoek geen effecten bij rokers gevonden van message framing of van tijdframe, noch op de attitude tegenover stoppen met roken, noch op de intentie om dat binnen drie maanden te gaan doen. Ook bij niet-rokers werden geen effectverschillen gevonden. De statistische

power van de toetsen in dit replicatieonderzoek maakte het voldoende mogelijk dat eventuele werkelijk bestaande effecten in dit experiment ook gevonden zouden worden. Dat was echter niet het geval.

Zoals in de inleiding vermeld, wijst eerder onderzoek op dit gebied overwegend een andere kant uit dan de uitkomsten die Mollen et al. (2017) rapporteren. Alles bijeen zijn de resultaten van het replicatieonderzoek dat wij uitvoerden veel meer in lijn met de in de inleiding genoemde eerdere studies.

Nadere bestudering van de uitkomsten van Mollen et al. (2017) leert ook dat hun bevindingen op ten minste één punt verrassend genoemd moeten worden: de hoge gemiddelde intentie-score van rokers die gain-framed boodschappen over kortetermijngevolgen van stoppen met roken te lezen kregen. De gemiddelde score ($M = 5.53$; $SD = 1.57$) lag hier ruim boven het schaal midden (4.0). In alle andere condities vonden Mollen et al. gemiddelden die lager waren dan het schaal midden. Wellicht heeft de toevallige samenstelling van de groep rokers in de gain-framed kortetermijnconditie hier een rol gespeeld. Anders is het moeilijk te verklaren dat rokende studenten na het zien van enkele anti-rookboodschappen zo gedecideerd het plan opvatten om met roken te stoppen, terwijl studenten die andere boodschappen te zien kregen dat in veel mindere mate deden. In het replicatieonderzoek was van een dergelijk effect op intentie ook niets te vinden. Rokende mbo-leerlingen die aan ons onderzoek meededen, scoorden in alle vier condities gemiddeld duidelijk lager op intentie om met roken te stoppen (tussen 2.35 en 2.51) dan het schaal midden (bij ons 3.0).

Hoe vallen de verschillen in uitkomsten van het oorspronkelijke onderzoek en het replicatieonderzoek te verklaren? Mogelijk zijn leeftijd en opleiding van belang: onze deelnemers waren gemiddeld jonger en hadden een lager opleidingsniveau dan de deelnemers aan het onderzoek van Mollen et al. (2017). Mogelijk ook speelden verschillen in het materiaal een rol: in het oorspronkelijke onderzoek werden alleen verbale waarschuwingen gebruikt, in ons onderzoek werden de teksten steeds vergezeld van een visualisatie. Ook weken we op één punt af van de thema's waar de boodschappen over gingen: terwijl in het oorspronkelijke onderzoek een deel van de boodschappen ging over effecten van roken op het uithoudingsvermogen, gebruikten wij in die gevallen boodschappen over effecten van roken op de huid.

Deze verschillen laten ons inziens de conclusie onverlet dat de empirische basis voor het advies van Mollen et al. (2017) aan de Europese Commissie niet stevig genoeg is. Zij raden aan om de nadruk leggen op kortetermijnvoordelen van stoppen met roken. Die lijken volgens hen in

elk geval bij jongeren die hoog zijn opgeleid het meest effectief. Dat mag zo zijn, maar bij lager opgeleide jongeren hebben wij voor zo'n advies in dit onderzoek geen steun gevonden. Wij zien ook geen aanleiding om bij de verkoop van sigarettenpakjes rekening te houden met het opleidingsniveau van de koper. Dat zou sowieso in de praktijk moeilijk uitvoerbaar zijn. Zinvoller is het om het onderzoek voort te zetten naar boodschappen die een zo breed mogelijke groep van rokers ertoe brengen om een punt achter hun rookgedrag te zetten.

Noten

- * Graag danken we Saar Mollen en haar collega's voor de informatie die zij ons gaven om dit replicatieonderzoek mogelijk te maken.
- 1. Alle powerberekeningen waarvan we hier verslag doen, zijn uitgevoerd met G*Power (Faul, Erdfelder, Lang, & Buchner, 2007; <http://www.gpower.hhu.de>)
- 2. We gebruikten de *Process* module van Hayes (2013).

Bibliografie

- Burgers, C., & Veldhuis, J. (2013). Tailoring fear appeals to lower-educated adolescents: The influence of modality and type of threat. *Information Design Journal*, 20(1), 32-46.
- Cauberge, V., De Pelsmacker, P., Janssens, W., & Dens, N. (2009). Fear, threat and efficacy in threat appeals: Message involvement as a key mediator to message acceptance. *Accident Analysis and Prevention*, 41(2), 276-285.
- DeCastellarnau, A. (2018). A classification of response scale characteristics that affect data quality: A literature review. *Quality & Quantity*, 52, 1523-1559.
- Eerland, A., & Van den Bergh, H. (2016). Empirische basis van conclusies: Handvatten voor de empirische taalonderzoeker. *Tijdschrift voor Taalbeheersing*, 38(2), 139-146.
- Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, 39(2), 175-191.
- Field, A. P. (2009). *Discovering Statistics Using SPSS*. (3rd, rev. ed.). London: SAGE.
- Gallagher, K. M., & Updegraff, J. A. (2012). Health message framing effects on attitudes, intentions, and behavior: A meta-analytic review. *Annals of Behavioral Medicine*, 43(1), 101-116.
- Green, L., & Myerson, J. (2004). A discounting framework for choice with delayed and probabilistic rewards. *Psychological Bulletin*, 130(5), 769-92.
- Hayes, A. F. (2009). Beyond Baron and Kenny: Statistical mediation analysis in the new millennium. *Communication Monographs*, 76(4), 408-420.
- Hayes, A. F. (2013). *Introduction to Mediation, Moderation, and Conditional Process Analysis: A Regression-Based Approach*. New York, NY: Guilford Publications.
- Hertgers, S., & Jansen, C. J. M. (2006). Wel vet, niet cool! De rol van consequentiebetrokkenheid bij de verwerking en het resultaat van fear appeal boodschappen in de obesitasvoorlichting. In B. Hendriks, H. Hoeken & P.J. Schellens (Red.), *Studies in taalbeheersing 2* (pp. 85-97). Assen: Van Gorcum.

- Hoeken, H., Hornikx, J., & Hustinx, L. (2012). *Overtuigende teksten: Onderzoek en ontwerp* (2e, herz. dr.). Bussum: Coutinho.
- Kantar TNS (2017). *Rapport kennis, norm en gedrag rondom de schadelijkheid van roken en meerooken*. Zie <https://www.rijksoverheid.nl/onderwerpen/roken/documenten/rapporten/2017/05/03/rapport-kennis-norm-en-gedrag-rondom-de-schadelijkheid-van-meerooken>
- Mollen, S., Engelen, S., Kessels, L. T. E., & Putte, B. van den (2017). Short and sweet: The persuasive effects of message framing and temporal context in antismoking warning labels. *Journal of Health Communication*, 22(1), 20-28.
- Nationaal Expertisecentrum Tabaksontmoediging, Onderdeel van het Trimbos-instituut (2017). *Factsheet roken onder volwassenen*. Zie <https://assets.trimbos.nl/docs/a264fcf9-a3e5-44c2-9ba6-e73cebd5d2ae.pdf>
- Noar, S. M., Hall, M. G., Francis, D. B., Ribisl, K. M., Pepper, J. K., & Brewer, N. T. (2016). Pictorial cigarette pack warnings: A meta-analysis of experimental studies. *Tobacco Control*, 25(3): 341-54.
- O'Keefe, D. J. (2003). Message properties, mediating states, and manipulation checks: Claims, evidence, and data analysis in experimental persuasive message effects research. *Communication Theory*, 13(3), 251-274.
- O'Keefe, D. J. (2016). *Persuasion: Theory and Research* (3rd ed.). Los Angeles, CA: Sage.
- O'Keefe, D. J., & Jensen, J. D. (2007). The relative persuasiveness of gain-framed loss-framed messages for encouraging disease prevention behaviors: A meta-analytic review. *Journal of Health Communication: International Perspectives*, 12(7), 623-644.
- O'Keefe, D. J., & Jensen, J. D. (2009). The relative persuasiveness of gain-framed loss-framed messages for encouraging disease detection behaviors: A meta-analytic review. *Journal of Communication*, 59(7), 296-316.
- Roken ontmoedigen door nieuwe regels* (z.j.). Deel van website van de rijksoverheid onder verantwoordelijkheid van het Ministerie van Volksgezondheid, Welzijn en Sport. Zie <https://www.rijksoverheid.nl/onderwerpen/roken/roken-ontmoedigen>
- Smith, K. H., & Stutts, M. A. (2003). Effects of short-term cosmetic versus long-term health fear appeals in anti-smoking advertisements on the smoking behaviour of adolescents. *Journal of Consumer Behaviour*, 3(2), 155-177.
- Tao, C.-C., & Bucy, E. P. (2007). Conceptualizing media stimuli in experimental research: Psychological versus attribute-based definitions. *Human Communication Research*, 33(4), 397-426.
- Volksgezondheidszorg.info (2018). Website gemaakt door het RIVM in opdracht van het Ministerie van Volksgezondheid, Welzijn en Sport. Zie <https://www.volksgezondheidszorg.info/onderwerp/roken/cijfers-context/huidige-situatie-volwassenen#node-roken-naar-opleiding>
- Zhao, X., & Peterson, E. (2017). Effects of temporal framing on response to antismoking messages: the mediating role of perceived relevance. *Journal of Health Communication*, 22(1), 37-44.